

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2005, 22:10:47 ; Search time 158 Seconds
(without alignments)
4149.071 Million cell updates/sec

Title: US-10-712-629B-2
Perfect score: 1374
Sequence: 1 agattcggaacagcagca.....ggccgcaggaccagctgcgc 746

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2_1/USFTO_spool_p/US10712629/runat_25112005_143035_10984/app_query.fasta_1.903
-DB=A.Geneseq -QWMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=us10712629 @CNC 1.1 348 @runat_25112005_143035_10984 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.Geneseq 21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1231	89.6	460	2 AAR74738	Aar74738 Human ubi
2	1231	89.6	461	2 AAR52980	Aar52980 Human rec
3	1231	89.6	461	2 AAR97982	Aar97982 Human ste
4	1231	89.6	461	2 AAR98140	Aar98140 NER recep
5	1231	89.6	461	5 ABG31015	Abg31015 Human LX
6	1231	89.6	461	8 ADJ14182	Adj14182 Human liv
7	1231	89.6	461	8 ADP05729	Adp05729 Human nuc
8	1231	89.6	461	8 ADQ48625	Adq48625 Wild type
9	1231	89.6	461	9 AEB94510	Aeb94510 Human LXR

10	1227	89.3	461	8 ADL12926	Adl12926 Human ste
11	1226	89.2	461	2 AAR96234	Aar96234 Human foe
12	1223	89.0	460	2 AAR25034	Aar25034 Human ubi
13	1221	88.9	364	2 AAR96235	Aar96235 Steroid h
14	1215.5	88.5	459	9 ADX85628	Adx85628 Human LXR
15	1215.5	88.5	460	7 ABR82544	Abr82544 Human LXR
16	1215.5	88.5	460	7 ABR84561	Abr84561 Human liv
17	1215.5	88.5	460	7 ADC81597	Adc81597 Human LXR
18	1215.5	88.5	460	8 ADN12168	Adn12168 Human LXR
19	1215.5	88.5	460	8 ABM82160	Abm82160 Tumour-as
20	1215.5	88.5	460	9 ADM88281	Adm88281 Human nuc
21	1215.5	88.5	460	9 AEB11855	Aeb11855 Liver x r
22	1162	84.6	297	3 AAB54183	Aab54183 Human pan
23	1101	80.1	446	2 AAR94169	Aar94169 OR-1 orph
24	1098.5	79.9	446	5 ABG31014	Abg31014 Mouse LX
25	1098.5	79.9	446	5 ABB57053	Abb57053 Mouse lsc
26	1098.5	79.9	446	8 ADP05731	Adp05731 Mouse nuc
27	1095.5	79.7	446	2 AAR99736	Aar99736 Retinoid
28	1094	79.6	443	2 AAR74739	Aar74739 Rat ubi
29	1094	79.6	443	2 AAW25035	Aaw25035 Rat ubi
30	1008	73.4	208	8 ADQ48626	Adq48626 Recombina
31	1005	73.1	263	8 ADJ14186	Adj14186 Human liv
32	969	70.5	417	9 AEB09222	Aeb09222 Human liv
33	777	56.6	264	5 ABG70080	Abg70080 Human pre
34	732	53.3	371	6 ABU96721	Abu96721 Human nuc
35	721.5	52.5	445	8 ADP05735	Adp05735 Mouse nuc
36	719.5	52.4	445	3 AAY32374	Aay32374 Mouse CNR
37	718.5	52.3	445	7 ADE62174	Ade62174 Rat Prote
38	718.5	52.3	445	7 ADD46146	Add46146 Rat Prote
39	707.5	51.5	358	7 ADC37546	Adc37546 Human nuc
40	707.5	51.5	358	8 ADG18111	Adg18111 Human LXR
41	707.5	51.5	377	6 ABR54168	Abr54168 Human NOV
42	706.5	51.4	440	2 AAR33744	Aar33744 XR2_3/20
43	705.5	51.3	293	6 ABR60037	Abp60037 LXR alpha
44	705.5	51.3	297	9 ADY51560	Ady51560 Wild-type
45	705.5	51.3	553	6 ABR83627	Abr83627 SUMO-LXRA

ALIGNMENTS

RESULT 1
AAR74738
ID AAR74738 standard; protein; 460 AA.

XX AAR74738;
AC AAR74738;
XX 20-JAN-1996 (first entry)
DT Human ubiquitous nuclear receptor protein.
DE Human ubiquitous nuclear receptor; drug design; disease diagnosis; therapy.
KW Ubiquitous nuclear receptor; protein.
XX Homo sapiens.
OS WO9513373-A1.
XX 18-MAY-1995.
XX 08-NOV-1994; 94WO-US012883.
XX 10-NOV-1993; 93US-00152003.
XX (ARCH-) ARCH DEV CORP.
XX Liao S, Song C;
PI WPI; 1995-194092/25.
DR N-PSDB; AAQ88760.
XX New ubiquitous nuclear receptor - used to develop prods. for use in
PT diagnostic, drug design and therapeutic applications.
XX Claim 5; Page 140; 195pp; English.

QY 301 GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
 DB 269 AlaHisPheThrGluLeuAlaIleSerValGlnGluLeuValAspPheAlaLysGln 288
 QY 361 GTCCCTGGTTCTCGAGCTGGCGGGAGGACAGATCGCCCTCTCTGAAGGATCCACT 420
 DB 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
 QY 421 ATCGAGATCATGCTGTAGACAGACGCGCGCTACAACCACGAGACAGAGTGATCACC 480
 DB 309 IleGluLeuMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
 QY 481 TTCTTGAG-GACTTCACCTACAGCAAGGACGACTTCACCGCTGACGGCTGCGAGTGGAG 539
 DB 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
 QY 540 TTCATCAACCCCATCTTCAGATTCTCGGGGCGCATCGCGGCTGGCGCTGGACGAGCT 599
 DB 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
 QY 600 GAGTACGCGCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGGCCCAACGTCAGGAG 559
 DB 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
 QY 660 CCGGCGCGGCTGGAGCGGTTGACAGAGCCCTACGTGGAGCGCTGCTGTCTACACGCGC 719
 DB 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
 QY 720 ATCAAGAGCGCGACGACGAGCTGGCC 746
 DB 409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 3

AAR97982
 ID AAR97982 standard; protein; 461 AA.
 AC AAR97982;
 XX
 DT 25-SEP-1996 (first entry)
 XX
 DE Human steroid receptor NER.
 KW Steroid receptor; NER; TOFA; NGF potentiator; muscarinic agonist;
 KW dopamine D1 antagonist.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 87..154
 FT /note= "putative double zinc finger structure"
 XX
 PN WO9613519-A1.
 XX
 PD 09-MAY-1996.
 XX
 PF 24-OCT-1995; 95WO-US013924.
 XX
 PR 27-OCT-1994; 94US-00330283.
 XX
 PA (MERI) MERCK & CO INC.
 PA (MEDI-) MEDICAL COLLEGE PENNSYLVANIA.
 XX
 PI Friedman E, Holloway MK, Rodan GA, Rutledge SJ, Schmidt A;
 PI Vogel RL;
 XX
 DR WPI; 1996-239449/24.
 DR N-PSDB; AAT18996.
 XX
 PT New isolated human steroid receptor NER - used to identify cpds. which
 PT can act, e.g., as potentiators of NGF, muscarinic agonists or dopamine D1
 PT antagonists.
 XX

PS Claim 5; Page 54-57; 72pp; English.

CC Steroid receptor NER (AAR97982) is a novel member of the nuclear receptor
 CC family. Its amino acid sequence was deduced from a gene (AAT18996)
 CC isolated from a human osteosarcoma SAOS-2/B10 cell library. A COS stable
 CC expression system allows prodn. of recombinant NER. NER is useful for
 CC identifying cpds. which act e.g. as potentiators of NGF, muscarinic
 CC agonists or dopamine D1 antagonists. TOFA (5-tetradecyloxy)-2-furan
 CC carboxylic acid) was found to activate NER and to be a potent potentiator
 CC of other drugs

XX Sequence 461 AA;

Alignment Scores:

Pred. No.: 1-27e-92 Length: 461
 Score: 1231.00 Matches: 248
 Percent Similarity: 99.60% Conservative: 0
 Best Local Similarity: 99.60% Mismatches: 1
 Query Match: 89.59% Indels: 1
 DB: 2 Gaps: 0

US-10-712-629B-2 (1-746) x AAR97982 (1-461)

QY 1 AAGATTTCGAAACAGCAGCAGGAGTCAAGTCAAGTCCAGTCCAGTTCGCTGGGCGC 60
 DB 169 LysIleArgLysGlnGlnGlnGlnSerGlnSerGlnSerGlnSerProValGlyPro 188
 QY 61 CAGGCGCAGCAGCAGCTCAGCCTCTGGGCTGGGGCTTCCCTGGTGGATCTGAGGAGGC 120
 DB 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
 QY 121 AGCCAGGGCTCCGGGGAAGCGGAGGTGTCCAGCTAACACAGCGCTCAAGAACTAATGATC 180
 DB 209 SerGlnGlySerGlyGlyGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
 QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACGCTCTCTTCCAGCACGCCCAA 240
 DB 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
 QY 241 GTCAGCCCTGGCCCTGGCGCAGACCCCGAGTCCCGAGATGCCCGCCAGCAACGCTTT 300
 DB 249 ValThrProTyrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
 QY 301 GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
 DB 269 AlaHisPheThrGluLeuAlaIleSerValGlnGluLeuValAspPheAlaLysGln 288
 QY 361 GTCCCTGGTTCTCGAGCTGGCGGGAGGACGACTCCACCGTGCAGGCTCCTGAAGGATCCACT 420
 DB 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
 QY 421 ATCGAGATCATGCTGTAGACAGACGCGCGCTACAACCACGAGACAGAGTGATCACC 480
 DB 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
 QY 481 TTCTTGAG-GACTTCACCTACAGCAAGGACGACTTCACCGTGCAGGCTGCGAGTGGAG 539
 DB 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
 QY 540 TTCATCAACCCCATCTTCAGATTCTCGGGGCGCATCGCGGCTGGCGCTGGACGAGCT 599
 DB 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
 QY 600 GAGTACGCGCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGGCCCAACGTCAGGAG 659
 DB 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
 QY 660 CCGGCGCGGCTGGAGCGGTTGACAGAGCCCTACGTGGAGCGCTGCTGTCTACACGCGC 719
 DB 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
 QY 720 ATCAAGAGCGCGACGACGAGCTGGCC 746

DB:	8	Gaps:	0
US-10-712-629B-2 (1-746) x ADU14182 (1-461)			
QY	1	AAGATTCGGAACAGCAGCAGGAGTACACAGTCACAGTCGCGAGTCACCTGTGGGCGG	60
Db	169	LysileArglysgInGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro	188
QY	61	CAGGGCAGCAGCAGCTCAGCTCTGGGCTGGGGCTTCCCTCGTGATCTGAGGCAGGC	120
Db	189	GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly	208
QY	121	AGCCAGGGCTCCGGGAAGCGAGGTGTCCAGCTAACAGCGCTCAAGAACTAATGATC	180
Db	209	SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGlnLeuMetile	228
QY	181	CAGCAGTTGGTGGCGCCCACTGCAGTGCAACAAAGCTCTCTCCGACAGCCCAAA	240
Db	229	GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys	248
QY	241	GTCCAGCCCTGGCCCTGGCGCAGACCCCGAGTCCCGAGATCCCGCCAGCAACGCTTT	300
Db	249	ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe	268
QY	301	GCCCACTTCACGAGCTGGCCATCTCAGTCCAGGAGATCTGGACTTCGCTAAGCAA	360
Db	269	AlaHisPheThrGluLeuAlaIleIleSerValGlnGlnIleValAspPheAlaLysGln	288
QY	361	GTCCCTGGTTCTGCAGTGGCGGAGGAGCAGATCCGCTCTCGAAGGCATCCACT	420
Db	289	ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr	308
QY	421	ATCGAGATCATGCTGTAGACAGCCAGCGCTTACAAACCAGCAGACAGAGTGTATCAC	480
Db	309	IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr	328
QY	481	TTCTTGAG-GACTTCACCTACAGCAGGACGACTTCACCGTCGAGCGCTTCAGGTGGAG	539
Db	329	PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu	348
QY	540	TTTCATCAACCCATCTCGAGTCTCGCGGCATCGCGGCTGGCGCTGGCGCTGGAGCGCT	599
Db	349	PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla	368
QY	600	GAGTACGCCCTGTCTATCGCATCAACATCTTCTCGGCCGACCGCCCAACGTCGAGGAG	659
Db	369	GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu	388
QY	660	CCGGGCGCGTGGAGCGGTTGCAGCAGCCCTACGTGAGGCGCTGTCTCTACACGGCG	719
Db	389	ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuSerTyrThrArg	408
QY	720	ATCAGAGCGCGCAGGACCGCTCGC	746
Db	409	IleLysArgProGlnAspGlnLeuArg	417
RESULT 7			
ADP05729			
ID	ADP05729 standard; protein; 461 AA.		
XX			
AC	ADP05729;		
XX			
DT	26-AUG-2004 (first entry)		
XX			
DE	Human nuclear receptor protein SeqID103.		
XX			
KW	disease risk; disorder risk; mutation; polymorphism;		
KW	nuclear receptor protein; antibacterial; antithyroid; cardiovascular-Gen;		
KW	cytostatic; dermatological; eating-Disorders-Gen; gastrointestinal-Gen;		
KW	gynaecological; hepatotropic; immunosuppressive; muscular-Gen;		
KW	nephrotropic; osteopathic; virucide; adrenal gland; colon;		
KW	cardiovascular; intestine; kidney; liver; lung; muscular; ovary; blood;		
KW	prostate; skin; spleen; stomach; testes; thymus; thyroid; uterus;		

Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 368
 QY 600 GAGTACGCCCTGCTCATCGCATCAACATCTTCTCGCGCGACCGGCCCAACGTGACAGGAG 659
 Db 369 GluTyrAlaLeuLeuLeuAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
 QY 660 CCGGGCCGCGTGGAGCGTGTGCAGCAGCCCTACGTGGAGCGCTGTGCTCTACACGCCG 719
 Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
 QY 720 ATCAAGAGCGCGCAGACGAGCTGCCG 746
 Db 409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 9

AEB94510
 ID AEB94510 standard; protein; 461 AA.

XX AC AEB94510;

XX DT 06-OCT-2005 (first entry)

XX DE Human LXR protein - SEQ ID 12.

XX KW lipid metabolism; insulin resistance;
 genetically engineered microorganism; screening; LXR.

XX OS Homo sapiens.

XX PN JP2005204550-A.

XX PD 04-AUG-2005.

XX PF 21-JAN-2004; 2004JP-00013728.

XX PR 21-JAN-2004; 2004JP-00013728.

XX PA (YAMA) YAMANOUCHI PHARM CO LTD.

XX PI Ida M, Ogino A, Endo H, Nishijima S, Inoki Y;

XX DR WPI; 2005-546101/56.

XX DR N-PSDB; AEB94509.

XX PT Screening lipid metabolism or insulin resistance improvement drug, by
 contacting test substance with cell capable of controlling ABCA1
 expression, analyzing change in liver X receptor dependent
 transcriptional activity.

XX PS Disclosure; SEQ ID NO 12; 25pp; Japanese.

XX CC The invention comprises a method of screening a drug for improving lipid
 metabolism or insulin resistance. The method involves contacting a test
 substance with a cell transformed by an expression vector encoding an LXR
 transcription coupling factor capable of controlling ABCA1 expression,
 and analyzing the effect of test substance on change in liver X receptor
 dependent transcriptional activity. The method of the invention is useful
 for screening a drug for improving lipid metabolism or insulin
 resistance. The present amino acid sequence represents a human LXR
 protein.

XX SQ Sequence 461 AA;

Alignment Scores:

Pred. No.:	1-27e-92	Length:	461
Score:	1231.00	Matches:	248
Percent Similarity:	99.60%	Conservative:	0
Best Local Similarity:	99.60%	Mismatches:	1
Query Match:	89.59%	Indels:	1
DB:	9	Gaps:	0

US-10-712-629B-2 (1-746) x AEB94510 (1-461)

QY 1 AAGATTGGAACACAGCAGCAGGAGTCAAGTCAAGTCCAGTCACTGTGGGGCCG 60
 Db 169 LysIleArgLysGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
 QY 61 CAGGGCAGCAGCAGCAGCTCGCCCTCGGGCTTCCCTGGTGGATCTGAGGCAGGC 120
 Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
 QY 121 AGCCAGGGTCCGGGGAAGCGAGGGTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
 Db 209 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
 QY 181 CAGCAGTTGGTGGCGGCCCAACTGACGTGCAACAAACGCTCTTCTCCACACGCCCAAA 240
 Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
 QY 241 GTCAGCCCTGGCCCTCGGGCGCAGACCCCGAGATGCCCGCAGCAGCAACCTTT 300
 Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 268
 QY 301 GCCCACTTCACGAGCTGGCCCATCTCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
 Db 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
 QY 361 GTGCTGTGTTCTCGAGCTGGCGGAGGAGACCAGATCGCCCTCTGAGGCGATCCACT 420
 Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
 QY 421 ATCGAGATCATGCTGTAGAGACAGCGCGCTACAACACAGACAGACAGAGTGTATCACC 480
 Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
 QY 481 TTCTTGAG-GACTTCACCTACAGCAGACGACTTCCACCGTCAGCGCTGCAGGTGGAG 539
 Db 329 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 348
 QY 540 TTCATCAACCCCATCTTCAGTTCTCGCGGCGCATCGCGGCTGGCGCTGGACGAGCT 599
 Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAla 368
 QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGCGCCCAACGTGACAGGAG 659
 Db 369 GluTyrAlaLeuLeuLeuAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
 QY 660 CCGGGCCGCGTGGAGCGGTTCAGCAGCCCTACGTGGAGCGCTGTGCTCTACACGCCG 719
 Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
 QY 720 ATCAAGAGCGCGCAGGACGAGCTGCCG 746
 Db 409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 10

ADL12926

ID ADL12926 standard; protein; 461 AA.

XX AC ADL12926;

XX DT 06-MAY-2004 (first entry)

XX DE Human steroid-induced C3A liver cell protein #105.

XX KW Hepatotropic; Gene therapy; Wilson disease; liver disorder;
 steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.

XX OS Homo sapiens.

XX PN US6673549-B1.

XX PD 06-JAN-2004.

XX PF 12-OCT-2001; 2001US-00976594.

PR 12-OCT-2000; 2000US-0240409P.
XX (INCY-) INCYTE CORP.
PA Furness LM, Buchbinder JL;
XX WPI; 2004-068610/07.
DR
XX
XX Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX
XX Disclosure; SEQ ID NO 655; 141pp; English.
XX
XX The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a cDNA. The sample is from a subject with
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human protein which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 461 AA;
SQ
Alignment Scores:
Pred. No.: 2,7e-92 Length: 461
Score: 1227.00 Matches: 247
Percent Similarity: 93.20% Conservatives: 0
Best Local Similarity: 99.20% Mismatches: 2
Query Match: 89.30% Indels: 1
DB: Gaps: 0
US-10-712-629B-2 (1-746) x ADL12926 (1-461)
QY 1 AAGATTTCGGAACAGCAGCAGCAGGAGTCACAGTCACAGTCACGTCACCTCTGGGGCG 60
Db 169 Lys11eArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCTGGGGCTTCCCTGGTGGATCTGAGGAGGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCCGGGAGCGAGGCTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGlyGlyGlyValGlnLeuThrAlaAlaGlnGlnLeuMetile 228
QY 181 CAGCAGTTGGTGGGGCCCAACTCAGTGCACAAACAGCTCTCTTCGACAGCCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCAGCCCTGCGCCCTGGGCGACACCCCGATCCGAGATCCCGCCAGCAAGCTTT 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
QY 301 GCCCACTTCACGAGCTGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleSerValGlnGlnLeuValAspPheAlaLysGln 288
QY 361 GTGCTGTGTTCTCCTCAGCTGGCGGGAGGACAGATCGCCCTCTCGAAGCATCCACT 420

Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAepGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCGAGATCATGCTGTAGAGACAGCAGCGCTACAAACCACAGACAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCTTTGAG-GACTTCACCTACAGCAGGACGACTTCCACCGTCGAGCGCTGCAGGTGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTATCAACCCCATCTTCGAGTTCTCGCGGCATCGCGCGCTGGGCTGTCTACACGCCGC 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla 368
QY 600 GAGTACGCGCTGCTCATCGCCATCAACATCTTCTCGCGCAGCGCGCCACGTCAGGAG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCGCGCTGGAGCGCTTGAGCAGCAGCCCTACGTCGAGCGCGCTGTCTCTACACGCCGC 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluThrLeuLeuSerTyrThrArg 408
QY 720 ATCAAGAGCGCGCAGGACCGAGTGGCGC 746
Db 409 IleLysArgProGlnAspGlnLeuArg 417
RESULT 11
AAR96234
ID AAR96234 standard; protein; 461 AA.
XX AAR96234;
AC AAR96234;
XX
DT 06-NOV-1996 (first entry)
XX
DE Human foetal lung steroid hormone receptor analogue ECDN.
KW Human; foetal lung; steroid hormone; receptor; analogue protein;
KW ECDN protein; cancer; screening; binding molecule; recombinant;
KW identification; anticancer drug; cancerous tissue; primer; probe;
KW antibody; immunohistochemical assay.
OS Homo sapiens.
XX
PN WO9609324-A1.
XX
PD 28-MAR-1996.
XX
XX 21-SEP-1995; 95WO-JP001909.
XX
XX 21-SEP-1994; 94JP-00226270.
XX
XX (CANC-) CANCER INST.
XX (EISA) EISAI CO LTD.
XX
XX Nakamura Y, Saito H;
XX
XX WPI: 1996-188403/19.
XX N-PSDB; AAT27616.
XX
PT ECDN protein, a steroid hormone receptor analogue from human foetal lung
PT - is expressed in cancer cells and is useful for cancer diagnosis and
PT drug development.
XX
PS Claim 1; Page 21-25; 43pp; Japanese.
XX
CC The present sequence is the human foetal lung derived steroid hormone
CC receptor analogue protein, ECDN. In various cancer cells a variant ECDN
CC protein, designated ECDN small mol. (ECDNsm) protein, is expressed.
CC Therefore screening for ECDN and ECDNsm protein binding mols., using
CC recombinant ECDN and ECDNsm proteins will be useful in the identification
CC of candidate anticancer drugs. Gene expression of ECDN and ECDNsm
CC proteins in normal and cancerous tissues can be studied using primers and
CC probes derived from ECDN and ECDNsm protein cDNA. Antibodies which

Db 288 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 307
QY 421 ATCGAGATCATGCTAGAGACAGCCAGCGCTACAAACCACGACAGAGTGTATCACC 480
Db 308 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 327
QY 481 TTCTTGAG-GACTTTCACCTACAGCAAGAGCAGCTTCCACCGTCAGGCGCTGCAGGTGGAG 539
Db 328 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 347
QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGCGCATGCGCGCTGGCGCTGGACGACGCT 599
Db 348 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 367
QY 600 GAGTACGCGCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGGCCACCGTGCAGGAG 659
Db 368 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 387
QY 660 CCGGGCGCGTGGAGCGTTGCAGCAGCCCTACGTGGAGCGCTGTGTCTTACACGCGC 719
Db 388 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuSerTyrThrArg 407
QY 720 ATCAAGAGCGCGACGACGCTGCGC 746
Db 408 TyrLysArgProGlnAspGlnLeuArg 416

RESULT 13

AAR96235
ID AAR96235 standard; protein; 364 AA.

AC AAR96235;

DT 06-NOV-1996 (first entry)

DE Steroid hormone receptor analogue ECDN small mol. variant.

KW Human; foetal lung; steroid hormone; receptor; analogue protein;
KW ECDN protein; cancer; screening; binding molecule; recombinant;
KW identification; anticancer drug; cancerous tissue; primer; probe;
KW antibody; immunohistochemical assay; variant; small molecule; ECDNsm.

OS Homo sapiens.

XX WO9609324-A1.

XX 28-MAR-1996.

XX 21-SEP-1995; 95WO-JP001909.

XX 21-SEP-1994; 94JP-00226270.

XX (CANC-) CANCER INST.

PA (EISA) EISAI CO LTD.

XX Nakamura Y, Saito H;

XX WPI; 1996-188403/19.

DR N-PSDB; AAT27617.

XX ECDN protein, a steroid hormone receptor analogue from human foetal lung
PT - is expressed in cancer cells and is useful for cancer diagnosis and
PT drug development.

PS Claim 3; Page 25-28; 43pp; Japanese.

XX The present sequence is the variant of the human foetal lung derived
CC steroid hormone receptor analogue protein ECDN, designated ECDN small
CC mol. (ECDNsm) protein. ECDNsm protein is expressed in various cancer
CC cells, therefore screening for ECDNsm protein binding moles., using
CC recombinant ECDNsm proteins will be useful in the identification of
CC candidate anticancer drugs. Gene expression of ECDNsm proteins in
CC cancerous tissues can be studied using primers and probes derived from

CC ECDNsm protein cDNA. Antibodies which recognise ECDNsm proteins can be
CC used in ECDNsm protein immunohistochemical assays
XX Sequence 364 AA;

Alignment Scores:

Pred. No.: 8.08e-92 Length: 364
Score: 1221.00 Matches: 246
Percent Similarity: 98.80% Conservative: 0
Best Local Similarity: 98.80% Mismatches: 3
Query Match: 98.86% Indels: 1
DB: Gaps: 2

US-10-712-629B-2 (1-746) x AAR96235 (1-364)

QY 1 AAGATTCCGAAACAGCAGCAGGAGTCACTCAGTCGCGAGTCACCTGTGGGGCGG 60
Db 72 LysIleArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 91

QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCGCTTCCCTCTGTGATCTGAGGACGCG 120
Db 92 GlnGlySerSerSerSerSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 111

QY 121 AGCCAGCGCTCCGGGAAGCGAGGTGTCCAGCTAACACGCGCTCAAGAACTAATGATC 180
Db 112 SerGlnGlySerGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 131

QY 181 CAGCAGTTGGTGGCGCGCCAACTGCAGTGCACAAACGCTCTTCCGACACGCGCAAA 240
Db 132 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 151

QY 241 GTCACGCGCTGGCGCGCGCAGACCCCGAGTCCCGAGATCCCGCCAGCAACGCTTT 300
Db 152 ValThrProThrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 171

QY 301 GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 172 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 191

QY 361 GTGCGTGGTTCTCGAGCTGGCGCGGAGGACCATCGCCCTCTCGAAGGATCCACT 420
Db 192 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleProLeuLeuAsnAlaSerThr 211

QY 421 ATCGAGATCATGCTAGACAGCAGCGCTACAAACCACGACAGAGTGTATCACC 480
Db 212 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 231

QY 481 TTCTTGAG-GACTTTCACCTACAGCAAGAGCAGCTTCCACCGTCAGGCGCTGCAGGTGGAG 539
Db 232 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 251

QY 540 TTCATCAACCCCATCTTCAGTTCTCGGGGCCATCGCGGCGCTGGCGCTGGACGCGCT 599
Db 252 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 271

QY 600 GAGTACGCGCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGCGCCACGTCGAGGAG 659
Db 272 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 291

QY 660 CCGGGCGCGTGGAGCGGTTGCAGCAGCGCTTACGTGGAGGCGCTGTGTCTTACACGCGC 719
Db 292 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuSerTyrThrArg 311

QY 720 ATCAAGAGCGCGCAGGACCGCTGCGC 746

Db 312 IleLysArgProGlnAspGlnLeuArg 320

RESULT 14

ADX85628

ID ADX85628 standard; protein; 459 AA.

XX

AC ADX85628;

XX

DT 05-MAY-2005 (first entry)
 XX Human LXRbeta protein.
 XX inflammatory bowel disease; LXRbeta; nuclear hormone receptor;
 KW metabolism; cholesterol; pharmaceutical; Crohn's disease;
 KW ulcerative colitis; colitis; ischemia; gastrointestinal-gen.;
 KW antiinflammatory; antiulcer; receptor.
 XX
 OS Homo sapiens.
 XX
 PN WO2005013946-A2.
 XX
 PD 17-FEB-2005.
 XX
 PF 27-JUL-2004; 2004WO-EP008426.
 XX
 PR 28-JUL-2003; 2003US-0490614P.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Goto Y, Kikkawa H, Kinoshita M;
 XX
 DR WPI; 2005-172945/18.
 DR N-PSDB; ADX85627.
 XX
 PT Preventing and/or treating inflammatory bowel diseases, such as Crohn's
 PT disease, ulcerative colitis and inflammatory colitis, comprises
 PT administering LXR agonist.
 PT
 PS Disclosure; SEQ ID NO 4; 55pp; English.
 XX
 CC The invention relates to a method of treating or preventing inflammatory
 CC bowel disease (IBD) in a mammal. The method comprises administering a LXR
 CC agonist, or its pharmaceutical salt, solvent, or physiologically
 CC functional derivative. The LXR agonist is a compound of formula (I) or
 CC (II), as defined in the specification. The LXR agonist is defined as any
 CC compound that enhances the biological activity of LXRalpha and/or
 CC LXRbeta. LXRalpha and LXRbeta (collectively known as LXR) are nuclear
 CC hormone receptors that regulate the metabolism of lipids such as
 CC cholesterol. The LXR agonists are useful in a pharmaceutical composition
 CC for preventing and/or treating inflammatory bowel diseases, such as
 CC Crohn's disease, ulcerative colitis, and inflammatory colitis caused by
 CC bacteria, ischemia, radiation, drugs or chemical substances. This
 CC sequence represents human LXRbeta protein.
 XX
 SQ Sequence 459 AA;
 Alignment Scores:
 Pred. No.: 2,39e-91 Length: 459
 Score: 1215.50 Matches: 247
 Percent Similarity: 99.20% Conservative: 0
 Best Local Similarity: 99.20% Mismatches: 1
 Query Match: 88.46% Indels: 2
 DB: 9 Gaps: 1
 US-10-712-629b-2 (1-746) x ADX85628 (1-459)
 QY 1 AAGATTCGGAACAGCAGCAGCAGGAGTCCAGTCACAGTCGAGTCACCTGTGGGGCGG 60
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 169 LysIleArgLys---GlnGlnGlnGlnSerGlnSerGlnSerGlnSerProValGlyPro 187
 QY 61 CAGGGCAGCAGCAGCTCAGCTTGGGCCCTTGGGCCCTTCCCTGGTGATCTGAGGAGGC 120
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 188 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyAlaSerProGlyGlySerGluAlaGly 207
 QY 121 AGCCAGGCTCCGGGAGCGAGGCTCCAGCTACAGCGCTCAAGAACTAATCATC 180
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 208 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetile 227
 QY 181 CAGCAGTTGGTGGCGGCCAACCTCAGTCGAAACAGCGCTCTTCTCCGACCCAGCCAAA 240
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 228 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 247

QY 241 GTCCACGCCCTGGCCCTGGGGCGAGACCCCGAGTCCCGAGATGCCCGCAGCAACGCTTT 300
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 248 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 267
 QY 301 GCCCACTTCACGAGCTGGCCCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 268 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 287
 QY 361 GTGCTTGGTTTCTCTGAGCTGGGCGGAGGACCATCGCCCTCTCTGAAGGATCCACT 420
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 288 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 307
 QY 421 ATCGAGATCATGCTGTAGACAGCCAGCGCTACAAACCGAGACAGAGTGTATCACC 480
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 308 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 327
 QY 481 TTCTTTCGAG-GACTTTCACCTACAGCAGGACGACTTCCACGCTGCAGCGCTGCAGGTGGAG 539
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 328 PheLeuLysAspPheThrTyrSerLysAspPhePheHisArgAlaGlyLeuGlnValGlu 347
 QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGCCCATCGCGGCTGGGCTTGGACGACGCT 599
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 348 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 367
 QY 600 GAGTACGCCCTGCTCATCGCCCATCAACATCTTCTCGCGCGACCGCCCAACGTCGAGGAG 659
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 368 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 387
 QY 660 CCGGGCGCGGTGGAGCGCTTGCAGCAGCCCTACGTGGAGCGCTGTGTCTCTACACGCGC 719
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 388 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 407
 QY 720 ATCAAGAGCGCGCAGACCCAGCTGCGC 746
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 408 IleLysArgProGlnAspGlnLeuArg 416
 RESULT 15
 ID ABR82544
 AC ABR82544;
 XX
 XX 20-NOV-2003 (first entry)
 DT
 XX
 DE Human LXR beta polypeptide.
 XX
 KW Pyridone; antilipemic; antidiabetic; antiarteriosclerotic; litholytic;
 KW hepatotropic; antiseborrheic; dermatological; antiparkinsonian; cardiant;
 KW nootropic; neuroprotective; antiinflammatory; cytostatic; anorectic;
 KW cardiovascular; nephrotropic; antithyroid; ophthalmological; hypotensive;
 KW cerebroprotective; immunosuppressive; nuclear receptor; liver X receptor;
 XX LXR beta; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003059884-A1.
 XX
 PD 24-JUL-2003.
 XX
 PF 20-DEC-2002; 2002WO-US041306.
 XX
 PR 21-DEC-2001; 2001US-0342707P.
 XX
 PA (XCEP-) X-CEPTOR THERAPEUTICS INC.
 XX
 PI Bayne CD, Johnson AT, Lu S, Mohan R, Griffith RC;
 XX
 DR WPI; 2003-646029/61.
 DR N-PSDB; ACF36088.
 XX
 PT New N-substituted pyridone compounds useful for the treatment of e.g.
 PT hypercholesterolemia, hyperlipoproteinemia, hyperglycemia, diabetes

PT mellitus, dyslipidemia or atherosclerosis.

PS Disclosure; Page 517-519; 544pp; English.

XX
XX
XX
CC The invention relates to N-substituted pyridone compounds of specified formula. The compounds act as modulators of nuclear receptor, orphan CC nuclear receptors and liver X receptor (LXRalpha or LXRbeta). They are CC useful for the treatment, prevention or amelioration of symptoms of CC nuclear receptor mediated diseases or disorders which are modulated or CC affected by nuclear receptor activity or in which nuclear receptor CC activity is implicated e.g. hypercholesterolemia, hyperlipoproteinemia, CC hypertriglyceridemia, lipodystrophy, hyperglycemia, diabetes mellitus, CC dyslipidemia, atherosclerosis, gallstone disease, acne vulgaris, CC acneiform skin conditions, diabetes, Parkinson's disease, Alzheimer's CC disease, cancer, inflammation, immunological disorders, lipid disorders, CC obesity, conditions characterized by a perturbed epidermal barrier CC function, conditions of disturbed differentiation or excess proliferation CC of the epidermis or mucous membrane, cardiovascular disorders, symptoms CC of a disease or disorder affected by cholesterol, triglyceride or bile CC acid levels, symptoms of hypocholesterolemia and in vitro method for CC altering nuclear receptor activity. The compounds are also used to reduce CC cholesterol levels, modulates cholesterol metabolism, increase CC cholesterol efflux from cells and increase the expression of ATP-Binding CC Cassette (ABC1) in the cells. The compounds are also useful for the CC treatment of macrovascular and macrovascular diseases e.g. nephropathy, CC neuropathy, retinopathy, hypertension, cerebrovascular disease, coronary CC heart disease, type 1 diabetes, type 2 diabetes, hyperlipidemia, Cushing CC syndrome, hypothyroidism, renal failure. The present sequence represents CC a human LXR beta polypeptide

XX
SQ Sequence 460 AA;

Alignment Scores:
Pred. No.: 2,39e-91 Length: 460
Score: 1215.50 Matches: 247
Percent Similarity: 99.20% Conservative: 0
Best Local Similarity: 99.20% Mismatches: 1
Query Match: 88.46% Indels: 2
DB: Gaps: 1

US-10-712-629b-2 (1-746) x ABR82544 (1-460)

QY 1 AAGATTGGAAACAGCAGCAGGAGTCACTGCGAGTCACTGCGGCGG 60
DB 169 LysileArgLys---GlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 187
QY 61 CAGGGCAGCAGCAGCTCAGCTCGGGCTGGGGCTTCCCTGGTGATCTGAGGAGGC 120
DB 188 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 207
QY 121 AGCCAGGGCTCCGGGAGCGAGGGGTGTCAGCTAACAGCGGCTCAAGAACTAATGATC 180
DB 208 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnLeuMetile 227
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACGGCTCTTCCGACAGCCCAAA 240
DB 228 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 247
QY 241 GTCACGCCCTGGGCCCTGGGGCGAGACCCCGAGTCCCGAGATGCCCGCAGCAACGCTTT 300
DB 248 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 267
QY 301 GCCCACTTCAGCGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
DB 268 AlaHisPheThrGlnLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 287
QY 361 GTGCTCTGGTTTCTCTGAGTGGCGGAGGACAGATCGCCCTCTCTGAAGGCATCCACT 420
DB 288 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 307
QY 421 ATCGAGATCATGCTGAGAGCAGCGCGCTACAACACAGACAGAGTGATATCACC 480
DB 308 IleGluIleMetLeuLeuGluThrAlaArgArgTyArgAsnHisGluThrGluCysIleThr 327

QY 481 TTCTTGAG-GACTTCACCTACAGCAAGGACGACTTCCACCGTGCAGSCCTGCAGGTGGAG 539
DB 328 PheLeuLysAspPheThrTySerLysAspPheHisArgAlaGlyLeuGlnValGlu 347
QY 540 TTATCAACCCCATTTCTCGAGTTCTCGCGGCCATCGCGGCTGGGCTGGACGACGCT 599
DB 348 PheIleAsnProIlePheGluPheSerArgAlaMetArgLysLeuGlyLeuAspAla 367
QY 600 GAGTAGCCCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGGCCCAACGTCAGGAG 659
DB 368 GluTyAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 387
QY 660 CCGGGCGCGTGGAGCGTTCGAGCAGCCCTACGTGGAGCGCTGTCTCTACACGCGC 719
DB 388 ProGlyArgValGluAlaLeuGlnProTyValGluAlaLeuLeuSerTyThrArg 407
QY 720 ATCAAGAGCCCGCAGGACCGAGCTGGCG 746
DB 408 IleLysArgProGlnAspGlnLeuArg 416

Search completed: November 25, 2005, 22:22:56
Job time : 172 secs

This Page Blank (Page 4)

A;Title: Isolation of proteins that interact specifically with the retinoid X receptor:
A;Reference number: A57664; MUID:95280959; PMID:7760852

A;Accession: I49021
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-446 <RES>
C;Genetics:
A;Gene: Rip15
C;Keywords: zinc finger
F;76-366/Domain: erba transforming protein homology <ERBA>

Alignment Scores:
Pred. No.: 2,76e-77 Length: 446
Score: 1098.50 Matches: 224
Percent Similarity: 93.57% Conservative: 9
Best Local Similarity: 89.96% Mismatches: 15
Query Match: 79.95% Indels: 2
DB: 2 Gaps: 1

US-10-712-629B-2 (1-746) x I49021 (1-446)

```
QY 1 AAGATTCGGAACAGCAGCAGGAGTCACTACAGTCAGTCGCGAGTCACCTGTGGGGCCG 60
   :::::::::::::::::::: ||||| ||||| |||||
Db 155 GlnileArglyslsArglleGlnlys---GlnGlnGlnGlnProProProSer 173
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCTGGGGCTTCCCTGCGATCTGAGGCAGGC 120
   :::::::::::::::::::: ||||| ||||| |||||
Db 174 GluProAlaAlerSerSerGlyArgProAlaAlerProGlyThrSerGluAlaSer 193
QY 121 AGCCAGGGTCCGGGAAGCGAGGTGTCAGCTAACAGCGGCTCAAGAACTAATGATC 180
   :::::::::::::::::::: ||||| ||||| |||||
Db 194 SerGlnGlySerGlyGluGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetile 213
QY 181 CAGCAGTTGGTGGCGGCCCACTGCGAGTGCACAAACAGCTCTTCTCCGACAGCCCAA 240
   :::::::::::::::::::: ||||| ||||| |||||
Db 214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTCACGCCCTGGGCCCTGGCGGAGACCCAGTCCGAGATGCCCGCCAGCAACGCTTT 300
   :::::::::::::::::::: ||||| ||||| |||||
Db 234 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 253
QY 301 GCCCACTTCACGAGCTGCCCATCTCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
   :::::::::::::::::::: ||||| ||||| |||||
Db 254 AlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLysGln 273
QY 361 GTGCCCTGGTTCTCTGAGCTGGCGGGAGGAGCAGATCGCCCTCTGAGGCGATCCACT 420
   :::::::::::::::::::: ||||| ||||| |||||
Db 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY 421 ATCGAGATCATGCTAGAGCAGCGCGCTACAAACACGAGACAGAGTGATCACC 480
   :::::::::::::::::::: ||||| ||||| |||||
Db 294 IleGluIleWetLeuLeuGluThrAlaArgArgAsnHisGluThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTACACAGCAGGAGCTTCCACCGTCGAGCGCTGCGAGTGGAG 539
   :::::::::::::::::::: ||||| ||||| |||||
Db 314 PheLeuLysAspPheThrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTCATCAACCCCNCTTCTCAGTTCTCGGGGCGATCGGGCGCTGGCGCTGAGCAGCGT 599
   :::::::::::::::::::: ||||| ||||| |||||
Db 334 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAla 353
QY 600 GAGTAGCCCTGTCTATCCCATCAACATCTTCTCGGGCGAGCCGCCCAACGTCGAGGAG 659
   :::::::::::::::::::: ||||| ||||| |||||
Db 354 GluTyAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY 660 CCGGGCGCGGTGGAGCGTTGACGACGCCCTACGTGGAGCGCTGCTGTCTACAGCGCG 719
   :::::::::::::::::::: ||||| ||||| |||||
Db 374 ProSerArgValGluAlaLeuGlnProTyValGluAlaLeuLeuSerTyThrArg 393
QY 720 ATCAGAGCGCGGAGGACGAGTGGCG 746
   :::::::::::::::::::: ||||| ||||| |||||
Db 394 IleLysArgProGlnAspGlnLeuArg 402
```

RESULT 4

A56043

steroid hormone receptor-like protein RLD-1 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 31-Dec-2004

C;Accession: A56043

R;Apfel, R.; Benbrook, D.; Lernhardt, E.; Ortiz, M.A.; Salbert, G.; Pfehl, M.

Mol. Cell. Biol. 14, 7025-7035, 1994

A;Title: A novel orphan receptor specific for a subset of thyroid hormone-responsive el

A;Reference number: A56043; MUID:95021230; PMID:7935418

A;Accession: A56043

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-445 <APP>

A;Cross-references: UNIPROT:Q62685; UNIPARC:UPI000013050D; CB:U11685; NID:9555751; PIDN

A;Note: authors translated the codon GAG for residue 73 as Ser

C;Keywords: steroid hormone receptor; zinc finger

F;94-365/Domain: erba transforming protein homology <ERBA>

Alignment Scores:

Pred. No.: 7,46e-48 Length: 445
Score: 718.50 Matches: 151
Percent Similarity: 75.40% Conservative: 36
Best Local Similarity: 60.89% Mismatches: 38
Query Match: 52.29% Indels: 24
DB: 2 Gaps: 2

US-10-712-629B-2 (1-746) x A56043 (1-445)

```
QY 1 AAGATTCGGAACAGCAGCAGGAGTCACTACAGTCAGTCGCGAGTCACCTGTGGGGCCG 60
   :::::::::::::::::::: ||||| ||||| |||||
Db 176 LysLeuLysArgGlnGlnGlnGlnAlaGlnAlaThrSerValSerPro----- 192
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCTGGGGCTTCCCTGCTGGATCTGAGGCAGGC 120
   :::::::::::::::::::: ||||| ||||| |||||
Db 193 -----ArgValSerSerProGlnValLeuPro----- 202
QY 121 AGCCAGGGTCCGGGAAGCGAGGTGTCAGCTAACAGCGGCTCAAGAACTAATGATC 180
   :::::::::::::::::::: ||||| ||||| |||||
Db 203 -----GlnLeuSerProGlnGlnLeuGlyMetile 212
QY 181 CAGCAGTTGGTGGCGGCCCACTGCGAGTGCACAAACAGCTCTTCTCCGACAGCCCAA 240
   :::::::::::::::::::: ||||| ||||| |||||
Db 213 GluLysLeuValAlaAlaGlnGlnGlnCysAsnArgArgSerPheSerAspArgLeuArg 232
QY 241 GTCACGCCCTGGGCCCTGGCGGAGACCCAGTCCGAGATGCCCGCCAGCAACGCTTT 300
   :::::::::::::::::::: ||||| ||||| |||||
Db 233 ValThrProTrpProIleAlaProAspProGlnSerArgGluAlaArgGlnArgPhe 252
QY 301 GCCCACTTCACGAGCTGCCCATCTCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
   :::::::::::::::::::: ||||| ||||| |||||
Db 253 AlaHisPheThrGluLeuAlaIleValSerValGlnGluIleValAspPheAlaLysGln 272
QY 361 GTGCCCTGGTTCTCTGAGCTGGCGGGAGGAGCAGATCGCCCTCTGAGGCGATCCACT 420
   :::::::::::::::::::: ||||| ||||| |||||
Db 273 LeuProGlyPheLeuGlnLeuSerArgGluAspGlnIleAlaLeuLeuLysThrSerAla 292
QY 421 ATCGAGATCATGCTCTAGAGACAGCCAGCGCTCAACACGAGACAGAGTGATCACC 480
   :::::::::::::::::::: ||||| ||||| |||||
Db 293 IleGluValMetLeuLeuGluThrSerArgArgTyArgAsnProGlySerGluSerIleThr 312
QY 481 TTCTTGAG-GACTTACCTACAGCAAGGAGCTTCCACCGTCGAGCGCTGCGAGTGGAG 539
   :::::::::::::::::::: ||||| ||||| |||||
Db 313 PheLeuLysAspPheSerTyArgGluAspPheAlaLysAlaGlyLeuGlnValGlu 332
QY 540 TTCATCAACCCCNCTTCTCAGTTCTCGGGGCGATCGGGCGCTGGCGCTGAGCAGCGT 599
   :::::::::::::::::::: ||||| ||||| |||||
Db 333 PheIleAsnProIlePheGluPheSerArgSerMetAsnGlnLeuGlnLeuAsnAspAla 352
QY 600 GAGTAGCCCTGTCTATCCCATCAACATCTTCTCGGGCGAGCCGCCCAACGTCGAGGAG 659
   :::::::::::::::::::: ||||| ||||| |||||
Db 353 GluPheAlaLeuLeuIleAlaIleSerIlePheSerAlaAspArgProAsnValGlnAsp 372
```

```
QY 660 CCGGGCCGCTGGAGCGTTTCAGCAGCCCTACGTGGAGCGCTGCTGCTTACACGCCG 719
Db ::::::::::::::::::::
373 GlnLeuGlnValGluArgLeuGlnHisThrTyrValGluAlaLeuHisAlaTyrValSer 392
QY 720 ATCAAGAGCGCCGAGCAGCAGCTG 743
Db ::::::::::::::::::::
393 IleAsnHisProHisAspArgLeu 400

RESULT 5
I38975
nuclear orphan receptor LXR-alpha - human
C:Species: Homo sapiens (man)
C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 05-Oct-2004
C:Accession: I38975
R:Willy, P.J.; Umesono, K.; Ong, E.S.; Evans, R.M.; Heyman, R.A.; Mangelsdorf, D.J.
Genes Dev. 9, 1033-1045, 1995
A:Title: LXR, a nuclear receptor that defines a distinct retinoid response pathway.
A:Reference number: I38975; MUID:95262897; PMID:7744246
A:Accession: I38975
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-447 <REP>
A:Cross-references: UNIPROT:Q13133; UNIPARC:UPI00000529F2; EMBL:U22662; NID:9726512; PID
C:Keywords: zinc finger
F:96-367/Domain: erba transforming protein homology <ERBA>

Alignment Scores:
Pred. No.: 1,29e-46 Length: 447
Score: 702.50 Matches: 149
Percent Similarity: 74.19% Conservative: 35
Best Local Similarity: 60.08% Mismatches: 41
Query Match: 51.13% Indels: 24
DB: 2 Gaps: 2

US-10-712-629B-2 (1-746) x I38975 (1-447)

QY 1 AAGATTCCGAAACAGCAGCAGCAGGAGTCACAGTCAGTCGAGTCACCTGTGGGCGG 60
Db ::::::::::::::::::::
178 LysLeuLysArgGlnGlnGluGlnGlnAlaHisAlaThrSerLeuProPro----- 194
QY 61 CAGGCGCAGCAGCTCAGCTCCCTGGGCTGGGGCTTCCCTCGTGGATCTGAGGCGG 120
Db ::::::::::::::::::::
195 -----ArgArgSerSerProGlnIleLeuPro----- 204
QY 121 AGCCAGGCTCCGGGAAGCGAGGTCTCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db ::::::::::::::::::::
205 -----GlnLeuSerProGlnGlnLeuGlyMetIle 214
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACGCTCTTCTCCGACCCAGCCCAA 240
Db ::::::::::::::::::::
215 GluLysLeuValAlaAlaGlnGlnGlnCysAsnArgArgSerPheSerAspArgLeuArg 234
QY 241 GTACGCGCTGGCCCTGGCGGCGAGACCCCGAGTCCCGAGATCCCGCCGACGACGCTTT 300
Db ::::::::::::::::::::
235 ValThrProTrpProMetAlaProAspProHisSerArgGluAlaArgGlnGlnArgPhe 254
QY 301 GCCCATTACGAGGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
Db ::::::::::::::::::::
255 AlaHisPheThrGluLeuAlaIleValSerValGlnGluLeuValAspPheAlaLysGln 274
QY 361 GTGCTGTTTCTGTCAGCTGGGCGGAGGAGCAGATCGCCCTCTGTAAGGCATCCACT 420
Db ::::::::::::::::::::
275 LeuProGlyPheLeuGlnLeuSerArgGluAspGlnIleAlaLeuLeuLysThrSerAla 294
QY 421 ATCGAGATCATGCTGTAGAGACGACCGCGCTTACACACGACGAGAGTGTATCACC 480
Db ::::::::::::::::::::
295 IleGluValMetLeuLeuGluThrSerArgArgTyrAsnProGlySerGluSerIleThr 314
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGCAGCTCCACCGCTCCAGCGCTGCGAGTGGAG 539
Db ::::::::::::::::::::
315 PheLeuLysAspPheSerTyrAsnArgGluAspPheAlaLysAlaGlyLeuGlnValGlu 334
```

```
QY 540 TTCATCAACCCCATCTTCAGTTCTCGCGGGCCATGCGCGGCTGGGCTGACGACGCT 599
Db ::::::::::::::::::::
335 PheIleAsnProIlePheGluPheSerArgAlaMetAsnGluLeuGlnLeuAsnAspAla 354
QY 600 GAGTACGCGCTCTCATCGCCATCAACATCTTCTCGCGGCGAGCCGCCCAACCTGGCAGG 659
Db ::::::::::::::::::::
355 GluPheAlaLeuLeuIleAlaIleSerIlePheSerAlaAspArgProAsnValGlnAsp 374
QY 660 CCGGGCCGCTGGAGCGTTTCAGCAGCCCTACGTGGAGCGCTGCTGCTTACACGCCG 719
Db ::::::::::::::::::::
375 GlnLeuGlnValGluArgLeuGlnHisThrTyrValGluAlaLeuHisAlaTyrValSer 394
QY 720 ATCAAGAGCGCCGAGCAGCAGCTG 743
Db ::::::::::::::::::::
395 IleHisHisProHisAspArgLeu 402

RESULT 6
A56590
ecysteroid receptor homolog cERH - midge (Chironomus tentans)
C:Species: Chironomus tentans
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 31-Dec-2004
C:Accession: A56590
R:Imhof, M.O.; Rusconi, S.; Lezli, M.
Insect Biochem. Mol. Biol. 23, 115-124, 1993
A:Title: Cloning of a Chironomus tentans cDNA encoding a protein (cERH) homologous to
A:Reference number: A56590; MUID:93250857; PMID:8485513
A:Accession: A56590
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-536 <IMH>
A:Cross-references: UNIPROT:P49882; UNIPARC:UPI0000050DEA; GB:S60739; NID:9385893; PIDN
C:Keywords: DNA binding; steroid hormone receptor; zinc finger
F:113-432/Domain: erba transforming protein homology <ERBA>

Alignment Scores:
Pred. No.: 5,13e-15 Length: 536
Score: 294.50 Matches: 74
Percent Similarity: 60.11% Conservative: 39
Best Local Similarity: 39.36% Mismatches: 70
Query Match: 21.43% Indels: 6
DB: 2 Gaps: 3

US-10-712-629B-2 (1-746) x A56590 (1-536)

QY 151 CAGCTAACACGCGCTCAAGAACTAATGATCCAGAGTTCGTGGCGGCCCAACTGCAGTGC 210
Db ::::::::::::::::::::
274 GlnLeuThrAlaAsnGlnValAlaValIleTyrLysLeuIleTrpTyrGlnAspGlyTyr 293
QY 211 AACAAACGCTCTTCTCCGACCGAGCCCAAGTCACGCCCTGGCCCTGGCGGCGAGACCCC 270
Db ::::::::::::::::::::
294 GluGlnProSerGluGluAspLeuLysArgIleThrThrGluLeuGluGluGluAsp 313
QY 271 CAGTCCCGAGATGCGCGCCGAGCAACGCTTTGCCACTTCACGAGCTGGCCATCATCTCA 330
Db ::::::::::::::::::::
314 GlnGluHisGluAla-----AsnPheArgTyrIleThrGluValThrIleLeuThr 330
QY 331 GTCACGAGATCGTGAGCTTCGCTAAGCAAGTGCCTGTTCTCGAGCTGGCGCCGGAG 390
Db ::::::::::::::::::::
331 ValGlnLeuIleValGluPheAlaLysGlyLeuProAlaPheIleLysIleProGlnGlu 350
QY 391 GACGAGATCGCCCTCTGAGGATCCACTATCGAGATCATGCTGTAGAGACAGCCAGG 450
Db ::::::::::::::::::::
351 AspGlnIleThrLeuLeuLysAlaCysSerSerGluValMetMetLeuArgMetAlaArg 370
QY 451 CGCTACAAACACGAGACAGAGTGTATCACCTCTTCGAGGACTTCACC-TACAGCAAGGAC 509
Db ::::::::::::::::::::
371 ArgTyrAspHisAspSerSerIleLeuPheAlaAsnThrAlaTyrThrLysGln 390
QY 510 GACTTCCACCGCTGAGCGCTGCGAGTTCATCAACCCCATCTCGAGTTCCTCGCGG 569
Db ::::::::::::::::::::
391 ThrTyrGlnLeuAlaGlyMetGlu--GluThrIleAspAspLeuLeuHisPheCysArg 409
```

QY 570 GCCATGCGCGGCTGGCGCTGAGACGCTGAGTACGCCCTGCTCATCGCCATCAACATC 629
Db 410 GlnMetTyrAlaLeuSerIleAspAsnValGluTyrAlaLeuLeuThrAlaIleValle 429
QY 630 TTCTCGGCCACCGCCCAACAGTCGACGAGCGCGCGCGCTGAGGCGTTGCGACGACCC 689
Db 430 PheSer---AspArgProGlyLeuGluLysAlaGluMetValAspIleIleGlnSerTyr 448
QY 690 TACGTGAGCGCGCTGCTCTCTAC 713
Db 449 TyrThrGluThrLeuLysValTyr 456
RESULT 7
JC7230
vitamin D receptor subtype b - Paralichthys olivaceus
C;Species: Paralichthys olivaceus
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 05-Oct-2004
C;Accession: JC7230
R;Suzuki, T.; Suzuki, N.; Srivastava, A.S.; Kurokawa, T.
Biochem. Biophys. Res. Commun. 270, 40-45, 2000
A;Title: Identification of cDNAs encoding two subtypes of vitamin D receptor in flounder
A;Reference number: JC7229
A;Accession: JC7230
A;Molecule type: mRNA
A;Residues: 1-425 <SUZ>
A;Cross-references: UNIPROT:Q9IB74; UNIPARC:UPI00000FCBB4; DDBJ:AB037674
A;Experimental source: intestine
C;Comment: This receptor is an important factor in calcium homeostasis and bone formation
C;Genetics:
A;Gene: vdr-b
C;Superfamily: Vitamin D3 receptor (VDR)
C;Keywords: bone; calcium transport; DNA binding; intestine; vitamin D
Alignment Scores:
Pred. No.: 1.96e-14 Length: 425
Score: 287.00 Matches: 76
Percent Similarity: 45.63% Conservative: 44
Best Local Similarity: 28.90% Mismatches: 91
Query Match: 20.89% Indels: 53
DB: 2 Gaps: 4

US-10-712-629B-2 (1-746) x JC7230 (1-425)

QY 112 GAGCAGCGCAGCGGCTCCGGGAGCGGCTGTCAGCTCAAGCGGCTCAAGAA 171
Db 114 GluAlaGlnArgGluAlaGluArgGluAlaArgProArgLeuThrAspGluGlnSer 133
QY 172 CTAATGATCAGCAGTTGGTGGCGGCCAACTGCAGTGCAACAAACGCTCTCTCCGAC 231
Db 134 GlnValIleAlaMetLeuValGluAlaHisHisLysThrTyrAspAspSerTyrSerAsp 153
QY 232 -----CAGCCCAAGTCACGCCCTGGCCCTG----- 258
Db 154 PheCysArgPheArgProValArgGluGlyProValThrArgSerAlaSerArgAla 173
QY 258 ----- 258
Db 174 AlaSerLeuHisSerLeuSerAspAlaSerSerAspSerPheSerHisSerProGluSer 193
QY 259 -----GCCGCAGAC 267
Db 194 ValAspThrLysValAsnPheAsnAsnLeuLeuMetMetTyrGlnGluGlnGlySerSer 213
QY 268 CCCAGTCCCGAGATGCCCGCCAGCAACGCTTTGCC-----CACTTCAGGAGCTG 318
Db 214 ProAspSerSerGluGluGluGlySerSerPheSerMetLeuProHisLeuAlaAspLeu 233
QY 319 GCCATCATCTCAGTCAGAGATCTCGGACTTCGCTAAGCAAGTCGCTGTTCTCTGCAG 378
Db 234 ValSerTyrSerIleGlnLysValIleGlyPheAlaLysMetIleProGlyPheArgGlu 253
QY 379 CTGGCGGGAGACGAGATCGCCCTCTCTGAAGGCATCCCATATCGAGATCATGTGCTA 438
Db 379 CTGGCGGGAGACGAGATCGCCCTCTCTGAAGGCATCCCATATCGAGATCATGTGCTA 438

Db 254 LeuThrAlaGluAspGlnIleAlaLeuLeuLysSerSerAlaIleGluValIleMetLeu 273
QY 439 GAGACAGCCAGCGCTACACACGACGACAGAGTGTATCACCTT----CTTGAGGACTTC 494
Db 274 ArgSerAsnGlnSerPheAsnLeuGluAspMetSerTrpSerCysGlyAlaProAspPhe 293
QY 495 ACCTCAGCAGCAAGCAGACTTCCACCGCTGACAGGCTGCAGGTGGAGTTTCATCAACCCCATC 554
Db 294 LysTyrGlnIleSerAspValThrLysAlaGlyHisThrLeuGluLeuGluProLeu 313
QY 555 TTCGAGTTCTCGCGGCCATGCGCGGCTGGGCTGACACGCTGAGTACGCCCTGCTC 614
Db 314 ValLysPheGlnValGlyLeuLysLeuAsnLeuGlnGluGluHisValMetLeu 333
QY 615 ATCGCCATCAACATCTTCGCGCGCGACCGCCCAACGTCGACGAGCGCGCGCTGCGAG 674
Db 334 MetAlaIleCysLeuLeuSerProAspArgProGlyValGlnAspHisAlaArgIleGlu 353
QY 675 GCCTTGACGACGCGCTACGTGGAGGCGCTCTCTCTACACGCGCATCAAGAGCGCCGAG 734
Db 354 AlaLeuGlnAspArgLeuSerGluThrLeuGlnAlaTyrIleGlnLeuHisHisProGly 373
QY 735 GACCAGCTG 743
Db 374 GlyArgLeu 376
RESULT 8
A41055
ecdysone receptor - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Dec-2004
C;Accession: A41055
R;Koelle, M.R.; Talbot, W.S.; Segraves, W.A.; Bender, M.T.; Cherbas, P.; Hogness, D.S.
Cell 67, 59-77, 1991
A;Title: The Drosophila Ecr gene encodes an ecdysone receptor, a new member of the ster
A;Reference number: A41055; MUID:92005697; PMID:1913620
A;Accession: A41055
A;Molecule type: mRNA
A;Residues: 1-878 <KOE>
A;Cross-references: UNIPROT:P34021; UNIPARC:UPI00001246E6; GB:M74078; NID:g157317; PIDN
A;Note: the authors also sequenced genomic exon boundaries
C;Comment: mRNA for this receptor is heavily expressed at the end of the third larval i
C;Genetics:
A;Gene: FlyBase:Ecr
A;Cross-references: FlyBase:FBgn0000546
A;Introns: 227/1; 398/3; 446/3; 497/3
C;Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation; z
F;262-572/Domain: ecrba transforming protein homology <ERBA>
Alignment Scores:
Pred. No.: 2.52e-14 Length: 878
Score: 285.50 Matches: 85
Percent Similarity: 48.51% Conservative: 45
Best Local Similarity: 31.72% Mismatches: 87
Query Match: 20.78% Indels: 52
DB: 2 Gaps: 9
US-10-712-629B-2 (1-746) x A41055 (1-878)
QY 10 AAACAGCAGCAGCAGGAGTCACAGTCACAGTCGACGTACCTGTGGCGCGCAGGCGCAGC 69
Db 347 LysLysAlaGlnLysGluLysAspLysMetThrThrSerProSerSerGlnHisGlyGly 366
QY 70 AGCAGCTCA---GCCTCTCGGCGCTGGGGCTTCCCTGGTGGATCTCAGGAGCAGCAGCAG 126
Db 367 AsnGlySerLeuAlaSerGly-----GlyGlyGln 376
QY 127 GCCTCGCGGGAAGCGGCGGCTGTCCAGCTAACAGCGGCTCAA----- 168
Db 377 AspPheValLysGluLeuLeuAspLeuMetThrCysGluProProGlnHisAlaThr 396
QY 169 ---GAACATAATGATCAGCAGCTTGGTGGCGGCCCAACAGTCAGTCAACAAACCTCTTC 225
Db 169 ---GAACATAATGATCAGCAGCTTGGTGGCGGCCCAACAGTCAGTCAACAAACCTCTTC 225

Db 397 IieProLeuLeuProAspGluIleLeuAla-----LysCysGlnAlaArgAsnIle 413
QY 226 TCCGACGACCCAAAGTCAGCCCTGGCCCTGGCGCA----- 264
Db 414 -----ProSerLeuThrTyrAsnGlnLeuAlaValIleTyrIlysLeuIleTrpTyr 430
QY 265 -----GACCCCAAGTCCCGAGATGCCCGCAG----- 291
Db 431 GlnAspGlyTyrGluGlnProSerGluGluAspLeuArgArgIleMetSerGlnProAsp 450
QY 292 -----CAACGCTTTGCCACTTCACGAGCTTCACGAGCTGGCCATCTCA 330
Db 451 GluAsnGluSerGlnThrAspValSerPheArgHisIleThrGluIleThrIleLeuThr 470
QY 331 GTCCAGGAGATCGTGACTTCGTAAAGCAAGTCCCTGGTTCTCGCAGCTGGCGCGGAG 390
Db 471 ValGlnLeuIleValGluPheAlaLysGlyLeuProAlaPheThrLysIleProGlnGlu 490
QY 391 GACCAGATCGCCCTCTCGAAGGCATCCACTATCGAGATCATGCTGTAGAGACGACGAG 450
Db 491 AspGlnIleThrLeuLeuLysAlaCysSerSerGluValMetMetLeuArgMetAlaArg 510
QY 451 CGGTACAACACAGACAGAGTGTATCACCTTCTTGAGGACTTC-ACCTACACAGGAC 509
Db 511 ArgTyrAspHisSerSerAspSerIlePhePheAlaAsnAsnArgSerTyrThrArgAsp 530
QY 510 GACTTCCACCGTGCAGCGCTCGAGTGGAGTTCATCAACCCCATCTTCGAGTTCTCGCGG 569
Db 531 SerTyrIysMetAlaGlyMet---AlaAspAsnIleGluAspLeuLeuHisPheCysArg 549
QY 570 GCCATCGCGCGCTGGCCCTGGACGCGTGTAGTACGCTGTCTCATCGCCATCAACATC 629
Db 550 GlnMetPheSerMetLysValAspAsnValGluTyrAlaLeuLeuThrAlaIleValIle 569
QY 630 TTCTCGGCCACCGGCCCAACGTGCAGGACCGCGCGCTGTGAGGCTGTGACGAGCCCC 689
Db 570 PheSer---AspArgProGlyLeuGluLysAlaGlnLeuValGluAlaIleGlnSerTyr 588
QY 690 TAGCTGGAGCGCTGCTGTCTCTAC 713
Db 589 TyrIleAspThrLeuArgIleTyr 596

RESULT 9

JC7229
vitamin D receptor subtype a - Paralichthys olivaceus
C;Species: Paralichthys olivaceus
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 05-Oct-2004
C;Accession: JC7229
R;Suzuki, T.; Suzuki, N.; Srivastava, A.S.; Kurokawa, T.
Biochem. Biophys. Res. Commun. 270, 40-45, 2000
A;Title: Identification of cDNAs encoding two subtypes of vitamin D receptor in flounder
A;Reference number: JC7229
A;Accession: JC7229
A;Molecule type: mRNA
A;Residues: 1-420 <SUZ>
A;Cross-references: UNIPROT:Q91B73; UNIPARC:UPI00000FC24F; DDBJ:AB037673
A;Experimental source: intestine
C;Comment: This receptor is an important factor in calcium homeostasis and bone formation
C;Genetics:
A;Gene: vdr-a
C;Superfamily: Vitamin D3 receptor (VDR)
C;Keywords: bone; calcium transport; DNA binding; hormone receptor; intestine; vitamin D

Alignment Scores:
Pred. No.: 3.12e-13 Length: 420
Score: 271.50 Matches: 76
Percent Similarity: 44.80% Conservative: 36
Best Local Similarity: 30.40% Mismatches: 85
Query Match: 19.76% Indels: 54
DB: 2 Gaps: 4

US-10-712-629B-2 (1-746) x JC7229 (1-420)

QY 151 CAGCTAACAGCGGCTCAAGAACTATGATCCAGCAGTGGTGGCGGCCCACTGCAGTGC 210
Db 123 ArgLeuAsnGluGluGlnAlaArgMetIleSerSerLeuValGluAlaHisIleThr 142
QY 211 AACAAACGCTCTCTTCGCGACGACCCCAAGTCACGCCCTGGCTGGCGCAGACCCC 270
Db 143 TyrAspAlaSerTyrSerAspPheSerArgPheArgPro---ProValArgGluGlyPro 161
QY 271 CAGTCCCGCA-----GATCCCCCAGCAA 294
Db 162 ValThrArgSerAlaSerArgAlaAlaSerLeuHisSerLeuSerAspAlaSerSerAsp 181
QY 295 CCGCTTT----- 300
Db 182 SerPheAsnHisSerProGluSerValAspThrLysMetAsnPheSerAsnLeuLeuMet 201
QY 300 ----- 300
Db 202 MetTyrGlnAspGlyAlaSerSerProAspSerSerGluGluAsnThrLysLeuSerMet 221
QY 301 ---GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGAGAGATCGTGGACTTCGCTAAG 357
Db 222 LeuProHisLeuAlaAspLeuValSerTyrSerIleGlnLysValIleGlyPheAlaLys 241
QY 358 CAAGTGCCTGGTTCTCGCAGCTGGCGCGGAGGACCATCGCTCCCTCGTGAAGGCATCC 417
Db 242 MetIleProGlyPheArgAspLeuThrAlaGluAspGlnIleAlaLeuLeuLysSerSer 261
QY 418 ACTATCGACATCATGCTGTAGACAGCGCGCTACAAACCCACAGACAGAGTGTATC 477
Db 262 AlaIleGluIleIleMetLeuArgSerAsnGlnSerPheSerLeuGluAspMetSerTrp 281
QY 478 ACCTTCTT---GAGGACTTCACACAGCAAGACAGACTTCACCGTCGAGCGCTGCAG 533
Db 282 SerCysGlyGlyProAspPheLysTyrCysIleAsnAspValThrLysAlaGlyHisThr 301
QY 534 GTGAGTTTCATCAACCCCATCTTCGAGTTCTCGCGGCCCATCGCGCGCTGGCCCTGGAC 593
Db 302 LeuGluLeuLeuGluProLeuValLysPheGlnValGlyLeuLysLeuAsnLeuHis 321
QY 594 GAGCGTGAGTACGCCCTGCTCATCGCCATCAACATCTCTCGCGCCACCGCGCCCAACGTG 653
Db 322 GluGluGluHisValLeuLeuMetGlyIleCysLeuLeuSerProAspArgProGlyVal 341
QY 654 CAGGACCGCGCGCGTGGAGCGCTTCAGCAGCCCTACGTGGAGCGCTGTGTCTCTAC 713
Db 342 GlnAspHisAlaArgValGluGlnLeuGlnAspArgLeuProGluAlaLeuGlnAlaTyr 361
QY 714 ACGCGCATCAAGAGCGCGCAGGACACGCTG 743
Db 362 IleArgIleAsnHisProGlyGlyArgLeu 371

RESULT 10

A31761

1,25-dihydroxyvitamin D-3 receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 05-Oct-2004
C;Accession: A31761; A31367
R;Burmeister, J.K.; Wiese, R.J.; Maeda, N.; Deluca, H.F.
Proc. Natl. Acad. Sci. U.S.A. 85, 9499-9502, 1988
A;Title: Structure and regulation of the rat 1,25-dihydroxyvitamin D-3 receptor.
A;Reference number: A31761; MUID:89071726; PMID:2849110
A;Accession: A31761
A;Molecule type: mRNA
A;Residues: 1-423 <BUR>
A;Cross-references: UNIPROT:P13053; UNIPARC:UPI00001382C3; GB:J04147; GB:J03630; NID:92
R;Burmeister, J.K.; Maeda, N.; Deluca, H.F.
Proc. Natl. Acad. Sci. U.S.A. 85, 1005-1009, 1988
A;Title: Isolation and expression of rat 1,25-dihydroxyvitamin D-3 receptor cDNA.
A;Reference number: A31367; MUID:88124963; PMID:2829212
A;Accession: A31367
A;Molecule type: mRNA
A;Residues: R', 58-423 <BU2>

A;Gene: GDB:VDR

A;Cross-references: GDB:120487; OMIM:601769; OMIM:277440

A;Map position: 12q12-12q14

C;Superfamily: Vitamin D3 receptor (VDR); erba transforming protein homology

C;Keywords: DNA binding; transcription regulation; zinc finger

F;22-341/Domain: erba transforming protein homology <BRBA>

F;24-44/Region: zinc finger

F;60-84/Region: zinc finger

Alignment Scores:

Pred. No.: 1-21e-11 Length: 427
Score: 251.00 Matches: 68
Percent Similarity: 49.11% Conservative: 42
Best Local Similarity: 30.36% Mismatches: 100
Query Match: 18.27% Indels: 15
DB: 2 Gaps: 5

US-10-712-629B-2 (1-746) x A28200 (1-427)

```
QY 64 GGCAGCAGCAGCTCAGCCTCTGGGCTGGGGCTTCCCTGGTATCTGAGCGAGCAGC 123
Db 162 GlyGlySerHisProSerArgProAsnSerArgHisThrProSerPheSerGlyAsp 181
QY 124 CAGGGCTCCGGGAAGCGGAGGTGTCCAGCTAACAGCGCTCAAGAACCTAATGATCCAG 183
Db 182 SerSerSerCysSerAspHisCys--IleThrSerSerAsp-----MetMetAsp 198
QY 184 CAGTTGGTGGCGCCCAACTGCAGTGCACAAAGCTCTCTCCAGCAGCCCAAGTC 243
Db 199 SerSerSerPheSerAsnLeuAspLeuSerGluGluAsp---SerAspAspProSerVal 217
QY 244 ACCCCCTGGGCCCTGGGCGCAGACCCCGAGTCCCGAGATCCCGCCAGCAACGCTTTGCC 303
Db 218 Thr-----LeuGluLeuSerGlnLeuSerMetLeuPro 228
QY 304 CACTTCAGGAGTGGCCATCATCTCAGTCCAGGAGTCTGGACTTCGCTAAGCAAGTG 363
Db 229 HisLeuAlaAspLeuValSerTyrSerIleGlnLysValIleGlyPheAlaLysMetIle 248
QY 364 CTGGTTCTCTGAGTGGCGGAGGACCATCGCTCTCTGAGGCGATCCACTATC 423
Db 249 ProGlyPheArgAspLeuThrSerGluAspGlnIleValLeuLeuLysSerSerAlaIle 268
QY 424 GAGATCATGCTGTAGAGACAGCAGCGCGCTACAAACACAGACAGAGTGTATCACTTC 483
Db 269 GluValIleMetLeuArgSerAsnGluSerPheThrMetAspAspMetSerTyrThrCys 288
QY 484 TT----GAGGACTTCACCTACAGCAAGGACGACTTCCACGCTGCAGGCTGCAGGTGGAG 539
Db 289 GlyAsnGlnAspTyrLysTyrArgValSerAspValThrLysAlaGlyHisSerLeuGlu 308
QY 540 TTCATCAACCCCATCTTCAGATTCTCGCGGCCATCGCGGGCTGGCGCTGGACGAGCT 593
Db 309 LeuIleGluProLeuIleLysPheGlnValGlyLeuLysLysLeuAsnLeuHisGluGlu 328
QY 600 GAGTACGCCCTGTCTATCGCCATCAACATCTTCTCGCGCCAGCCGCGCCCAACGTCAGGAG 659
Db 329 GluHisValLeuLeuMetAlaIleCysIleValSerProAspArgProGlyValGlnAsp 348
QY 660 CCGGGCGCGTGGAGCGGTGTGACAGACGCCCTACGTGGAGCGCTGTGTCCTACAGCGC 719
Db 349 AlaAlaLeuIleGluAlaIleGlnAspArgLeuSerAsnThrLeuGlnThrTyrIleArg 368
QY 720 ATCAAGAGCGCG 731
Db 369 CysArgHisPro 372
```

RESULT 15

A41977

retinoic acid receptor alpha - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 07-Feb-1997

C;Accession: A41977

R;Blumberg, B.; Mangelsdorf, D.J.; Dyck, J.A.; Bittner, D.A.; Evans, R.M.; De Robertis, Proc. Natl. Acad. Sci. U.S.A. 89, 2321-2325, 1992

A;Title: Multiple retinoid-responsive receptors in a single cell: families of retinoid

A;Reference number: A41977; PMID:92196110; PMID:1312717

A;Accession: A41977

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-459 <BLU>

A;Cross-references: UNIPARC:UPI0000177C1D

A;Experimental source: egg

A;Note: sequence extracted from NCBI backbone (NCBIP:88334)

C;Superfamily: retinoic acid receptor alpha; erba transforming protein homology

C;Keywords: zinc finger

F;86-337/Domain: erba transforming protein homology <ERBA>

F;88-108/Region: zinc finger

F;124-148/Region: zinc finger

Alignment Scores:

Pred. No.: 5-48e-11 Length: 459
Score: 242.50 Matches: 64
Percent Similarity: 54.55% Conservative: 38
Best Local Similarity: 34.22% Mismatches: 70
Query Match: 17.65% Indels: 16
DB: 2 Gaps: 4

US-10-712-629B-2 (1-746) x A41977 (1-459)

```
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAA-----CGCTCCTTCTCGGACCAG 234
Db 196 GlnGluThrPheProAlaLeuCysGlnLeuGlyLysTyrThrThrSerPheSerSerGlu 215
QY 235 CCCAAAGTCACGCCCTGGCGCCCGCCAGTCCCGGAGATGCCCGCAGATGCCCGCCAGCA 294
Db 216 GlnArgValSer-----LeuAspIleAspLeuTrpAspLys----- 227
QY 295 CGCTTTGCCCATTCACGAGGTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCT 354
Db 228 -----PheSerGluLeuSerThrLysCysIleIleLysThrValGluPheAla 243
QY 355 AGCAAGTGCCTGGTTTCTGCGAGCTGGCGGCGGAGCAGCAGATCGCCCTCCTGAAGGCA 414
Db 244 LysGlnLeuProGlyPheThrThrLeuThrIleAlaAspGlnIleThrLeuLeuLysSer 263
QY 415 TCACATCATCGAGATCATGCTGCTAGACAGCCAGCGCTACAAACACAGACAGAGTGT 474
Db 264 AlaCysLeuAspIleLeuIleLeuArgIleCysThrArgTyrThrProAspGlnAspThr 283
QY 475 ATCACCCTTTT-GAGGACTTTCACCTACAGCAAGGACGACTTCCACCGTCGAGGCTGCAG 533
Db 284 MetThrPheSerAspGlyLeuThrLeuAsnArgThrGlnMetHisAsnAlaGlyPheGly 303
QY 534 GTGGAGTTTCATCAACCCCATCTTCGAGTTCTCGCGGCCATCGCGGGCTGGCGCTGCAC 593
Db 304 ---ProLeuThrAspLeuValPheAlaPheAlaAsnGlnLeuLeuProLeuGluMetAsp 322
QY 594 GAGCTGAGTACGCCCTGTCTATCGCCATCAACATCTTCTCGCGCCAGCCGCCCAACGCTG 653
Db 323 AspAlaGluThrGlyLeuLeuSerAlaIleCysLeuIleCysGlyAspArgGlnAspLeu 342
QY 654 CAGGAGCCGGCGCGTGGAGCGTGTGAGGAGCCCTACGTGGAGGCGCTGTGTCTCTAC 713
Db 343 GluGlnProAspLysValAspLysLeuGlnLeuGluProLeuLeuGluAlaLeuLysIleTyr 362
QY 714 AGCGCATCAAGAGCGCGCAG 734
Db 363 ValArgThrArgArgProGln 369
```

Search completed: November 25, 2005, 22:31:41

Job time : 48 secs

the Reg Bank (uspio)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspro)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2005, 22:13:37 ; Search time 220 Seconds
(without alignments)
4784.765 Million cell updates/sec

Title: US-10-712-629B-2
Perfect score: 1374
Sequence: 1 aagattcgaaacagcagca.....ggccgcaggaccagctgcgc 746

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q/cgn2_1/USFTO.spool_p/US10712629/runat_25112005_143036_10996/app_query.fasta_1.903
-DB=UniProt -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10712629 @CNC 1.1 466 @runat_25112005_143036_10996 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -JONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1231	89.6	461	1 NRIH2_HUMAN	P55055 homo sapien
2	1231	89.6	461	2 Q510Y1_HUMAN	Q510Y1 homo sapien
3	1231	89.6	461	2 Q61BU6_HUMAN	Q61BU6 homo sapien
4	1189	86.5	460	2 Q68CY8_HUMAN	Q68CY8 homo sapien
5	1147.5	83.5	452	2 Q58D83_BOVIN	Q58D83 bos taurus
6	1144.5	83.3	455	1 NRIH2_BOVIN	Q58D83 bos taurus
7	1144.5	83.3	455	2 Q58D83_BOVIN	Q58D83 bos taurus
8	1144.5	83.3	498	2 Q58D83_BOVIN	Q58D83 bos taurus
9	1101	80.1	446	1 NRIH2_RAT	Q58D83 bos taurus
10	1098.5	79.9	446	1 NRIH2_MOUSE	Q58D83 bos taurus
11	1095.5	79.7	443	2 Q8BP65_MOUSE	Q8BP65 mus musculus
12	1028	74.8	351	2 Q5RE14_PONPY	Q5RE14 pongo pygma
13	855	63.0	441	2 Q6GNA2_XENLA	Q6GNA2 xenopus lae
14	731	53.2	412	2 Q5GA56_BRARE	Q5GA56 brachydanio
15	721.5	52.5	445	1 NRIH3_MOUSE	Q920Y9 mus musculus
16	719.5	52.4	445	2 Q91X41_MOUSE	Q91X41 mus musculus

17	719.5	52.4	447	2	Q4TU03_PIG	Q4TU03 sus scrofa
18	718.5	52.3	445	1	NRIH3_RAT	Q62685 rattus norv
19	718.5	52.3	445	2	Q51035_RAT	Q51035 rattus norv
20	712.5	51.9	447	2	Q5E9B6_BOVIN	Q5E9B6 bos taurus
21	707.5	51.5	402	2	Q81W13_HUMAN	Q81W13 homo sapien
22	704.5	51.3	447	1	NRIH3_HUMAN	Q13133 homo sapien
23	704.5	51.3	447	2	Q5U0N9_HUMAN	Q5U0N9 homo sapien
24	702.5	51.1	409	2	Q8JHUI_CHICK	Q8JHUI gallus gall
25	690.5	50.3	409	2	Q8AXU8_CHICK	Q8AXU8 gallus gall
26	499	36.3	509	2	Q4SHF7_TETNG	Q4SHF7 tetradodon n
27	472.5	34.4	557	2	Q4H386_CIOIN	Q4H386 ciona intes
28	400.5	29.1	444	2	Q44336_9ACAR	Q44336 amblyomma a
29	400.5	29.1	560	2	Q44337_9ACAR	Q44337 amblyomma a
30	400.5	29.1	570	2	Q44338_9ACAR	Q44338 amblyomma a
31	331.5	24.1	328	2	Q5GCP7_GECIA	Q5GCP7 secaracinus
32	331.5	24.1	518	2	Q76246_UCAPU	Q76246 uca pugilac
33	325.5	23.7	488	2	Q4W6C9_LEPDE	Q4W6C9 leptinotars
34	325.5	23.7	565	2	Q4W6D0_LEPDE	Q4W6D0 leptinotars
35	321	23.4	541	2	Q97095_LOCMI	Q97095 locusta mig
36	312.5	22.7	336	2	Q6RIB4_CARMA	Q6RIB4 carcinus ma
37	307	22.3	491	2	Q02035_TENMO	Q02035 tenebrio mo
38	299	21.8	503	2	Q6P7H5_XENLA	Q6P7H5 xenopus lae
39	299	21.8	513	2	Q8JU28_XENLA	Q8JU28 xenopus lae
40	297.5	21.7	530	2	Q8JU27_XENLA	Q8JU27 xenopus lae
41	294.5	21.4	536	1	ECR CHITE	P49882 chironomus
42	291.5	21.2	556	1	ECR MANSE	P49883 manduca sex
43	289.5	21.1	412	2	Q7PVB2_ANOGA	Q7PVB2 anopheles g
44	289.5	21.1	513	2	Q77240_CHOFU	Q77240 choristoneu
45	289.5	21.1	541	2	Q77255_CHOFU	Q77255 choristoneu

ALIGNMENTS

RESULT 1

NRIH2_HUMAN

ID NRIH2_HUMAN STANDARD; PRT; 461 AA.

AC P55055; Q12970;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Oxytetracycline receptor LXR-beta (Liver X receptor beta) (Nuclear orphan

receptor LXR-beta) (Ubiquitously-expressed nuclear receptor) (Nuclear

receptor LXR-beta) (NER, UNR;

GN Name=NRIH2; Synonyms=LXRB, NER, UNR;

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RP [1]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Osteosarcoma;

RX MEDLINE=95011628; PubMed=7926814; DOI=10.1016/0378-1119(94)90080-9;

RA Shinar D.M., Endo N., Rutledge S.J., Vogel R., Rodan G.A., Schmidt A.;

RT "NER, a new member of the gene family encoding the human steroid

hormone nuclear receptor.";

RL Gene 147:273-276(1994).

RN [2]

RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RP TISSUE=Blood, Placenta, and Uterus;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Shenmen C.M., Schuler G.D.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M.J., Udell T.B., Toshiyuki S., Carninci P., Schaefer T.E.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McWaters K.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,


```
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINUR.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; Znf C4; 1.
DR PROSITE; PS00031; NUCLEAR_REC_NBD 1; 1.
DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
KW SEQUENCE 461 AA; 51087 MW; 697E3D9F9BC5C0BE CRC64;

Alignment Scores:
Pred. No.: 4,24e-80 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservativeness: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 2 Gaps: 0

US-10-712-629B-2 (1-746) x Q61BU6_HUMAN (1-461)
QY 1 AGATTTCGGAACAGCAGCAGGAGTCACAGTCACAGTCGAGTCACCTGTGGGGCCG 60
Db 169 LysileArglysglnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCTCAGCTCTGGGCGTGGGGCTTCCCTCGTGATCTGAGGCGAGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGCGTCCGGGGAAGCGAGGTGTCCAGCTAACAGCGGTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyGlyValGlnLeuThrAlaGlnGluLeuMetIle 228
QY 181 CAGCAGTTGTGGGGCCCACTGCAGTGCAACAAAGCTCTCTTCCGACCCAGCCCAA 240
Db 229 GlnGlnLeuValAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCAGCGCCCTGGCGGCGAGACCCAGTCCGAGATCGCCGACGAGCAACGCTT 300
Db 249 ValThrProThrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnAlaGlyPhe 268
QY 301 GCCACTTCAGGAGTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTGCGTGGTTCTGAGCTGGCGGCGGAGACCATCGCTCCCTCTGAGGCGATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCAGATCATGCTGTAGACAGCAGCGGCTTACAAACACAGACAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCTTGAG-GACTTCACCTACAGCAGGACGACTTCCACCGTCGAGCGCTCGAGGTGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTCATCAACCCCATCTTCGAGTTCTCGGGGCCATGCGGGCTGGGGCTGGAGCAGCT 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
QY 600 GAGTACGCCCTGTCTATCGGCATCAACATCTTCTCGGCGGACCGGCCCAACGTCGAGGAG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCCGCTGGAGCGCTTGCAGCAGCCCTTACGTGGAGCGCTGTCTCTACACCGCC 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
QY 720 ATCAGAGGCGCGAGGACCGCTCGC 746
```

```
Db 409 IleLysArgProGlnAspGlnLeuArg 417
RESULT 4
ID Q68CY8_HUMAN PRELIMINARY; PRT; 460 AA.
AC Q68CY8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp686D1580.
GN Name=DKFZp686D1580;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]_
NUCLEOTIDE SEQUENCE.
TI TISSUE=Endometrium carcinoma cell line;
RG The German cDNA Consortium;
RA Ansoorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity). Receptor family.
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; CR749648; CAH18442.1; -; mRNA.
DR SMR; Q68CY8; 220-459.
DR Ensembl; ENSG00000131408; Homo sapiens.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000536; Hrmn_recept_lig.
DR InterPro; IPR001723; Strhmn_receptor.
DR InterPro; IPR000324; Vitd_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINUR.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; Znf C4; 1.
DR PROSITE; PS00031; NUCLEAR_REC_NBD 1; 1.
KW DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;
KW Receptor; Transcription; Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 460 AA; 50401 MW; FD0063D952B23D94 CRC64;

Alignment Scores:
Pred. No.: 4,45e-77 Length: 460
Score: 1189.00 Matches: 247
Percent Similarity: 98.41% Conservativeness: 0
Best Local Similarity: 98.41% Mismatches: 1
Query Match: 86.54% Indels: 4
DB: 2 Gaps: 0

US-10-712-629B-2 (1-746) x Q68CY8_HUMAN (1-460)
QY 1 AGATTTCGGAACAGCAGCAGGAGTCACAGTCACAGTCGAGTCACCTGTGGGGCCG 60
Db 169 LysileArglysglnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCTCAGCTCTGGGCGTGGGGCTTCCCTCGTGATCTGAGGCGAGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGCGTCCGGGGAAGCGAGGTGTCCAGCTAACAGCGGTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyGlyValGlnLeuThrAlaGlnGluLeuMetIle 228
```


US-10-712-629B-2 (1-746) x NR1H2_BOVIN (1-455)

QY 1 AGATTTCGGAACAGCAGCAGCAGCAGTCAAGTCCAGTGCAGTCCACTGTGGCGCG 60
| | | | |
Db 164 LysIleArgLysGlnGlnGln-----GlnGlnGlnSerProThrGlyPro 181

QY 61 CAGGGCAGCAGCAGCTCA---GCCTCTGGCGCTGGGGCTTCCCCTGTGGATCTCAGGCA 117
| | | | |
Db 182 GlyValSerSerSerProAlaSerGlyProGlyAlaSerProGlyGlySerAspGly 201

QY 118 GGCAGCCAGGGCTCCGGGAAGAGCGGTGTCAGCTAACAGCGGCTCAAGAATAATG 177
| | | | |
Db 202 GlyGlyGlnGlySerGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMet 221

QY 178 ATCCAGCAGTTGGTGGCGGCCAACTGCAGTGCAACAACGCTCTCTCTCCAGCACGCC 237
| | | | |
Db 222 IleGlnGlnLeuValAlaAalaGlnLeuGlnCysAenLysArgSerPheSerAspGlnPro 241

QY 238 AAAGTCACGCCCTGGCGCGCAGACCCTCCAGTCCCCGAGATGCCCGCAGCAACGC 299
| | | | |
Db 242 LysValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArg 261

QY 298 TTTGCCCACTTCACGAGCTGCCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAG 357
| | | | |
Db 262 PheAlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLys 281

QY 358 CAAGTGCCTGGTTTCTGCAGCTGGCGCGGAGGACACAGATCGCCCTCTCGAAGCATCC 417
| | | | |
Db 282 GlnValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSer 301

QY 418 ACTATCGAGATCATCTGCTAGAGACAGCCAGCGGCTACACACCAGACAGAGTGTATC 477
| | | | |
Db 302 ThrIleGluIleMetLeuLeuGluThrAlaArgArgTyrAenHisGluThrGluCysIle 321

QY 478 ACCTTCTTCAG-GACTTTCACCTACAGCAGGACGACTTCCACCGTCAGGCTGCAGGTG 536
| | | | |
Db 322 ThrPheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnVal 341

QY 537 GAGTTCATCAACCCCATCTTCGAGTTCTCGCGGGCCATCGCGCGCTGGGCTGCACAC 596
| | | | |
Db 342 GluPheIleAenProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAsp 361

QY 597 GCTGAGTAGCCCTGCTCATCCCATCAACATCTTCTCGCGGACCGCCCAACCTGCAG 656
| | | | |
Db 362 AlaGluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProasnValGln 381

QY 657 GAGCGCGCGCGCTGGAGCGTTTCAGCAGCGCCCTACGTGGAGCGCTGTGTCTACACG 716
| | | | |
Db 382 GluProSerArgValGluAlaLeuGlnProTyrValAspAlaLeuLeuSerTyrThr 401

QY 717 CGCATCAAGAGCCCGCAGCAGCAGCTGGCG 746
| | | | |
Db 402 ArgIleLysArgProGlnAspGlnLeuArg 411

RESULT 7

OSBINI_BOVIN PRELIMINARY; PRT; 455 AA.

ID OSBINI_BOVIN PRELIMINARY; PRT; 455 AA.
AC OSBINI;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DE Nuclear receptor subfamily 1, group H, member 2.
GN Name=NR1H2;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RX NUCLEOTIDE SEQUENCE.
RP TISSUE=Pooled;
RC MEDLINE=21180013; PubMed=11282978; DOI=10.1101/gr.170101;
RA Smith T.P.L., Grosse W.M., Freking B.A., Roberts A.J., Stone R.T.,

US-10-712-629B-2 (1-746) x NR1H2_BOVIN (1-455)

QY 1 AGATTTCGGAACAGCAGCAGCAGCAGTCAAGTCCAGTGCAGTCCACTGTGGCGCG 60
| | | | |
Db 164 LysIleArgLysGlnGlnGln-----GlnGlnGlnSerProThrGlyPro 181

QY 61 CAGGGCAGCAGCAGCTCA---GCCTCTGGCGCTGGGGCTTCCCCTGTGGATCTCAGGCA 117
| | | | |
Db 182 GlyValSerSerSerProAlaSerGlyProGlyAlaSerProGlyGlySerAspGly 201

QY 118 GGCAGCCAGGGTCCGGGAAAGCGAGGGTGTCAGCTAACAGCGGCTCAAGAATAATG 177
| | | | |
Db 202 GlyGlyGlnGlySerGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMet 221

QY 178 ATCCAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAACAGCTCTCTCTCCAGCACGCC 237
| | | | |
Db 222 IleGlnGlnLeuValAlaAalaGlnLeuGlnCysAenLysArgSerPheSerAspGlnPro 241

QY 238 AAAGTCACGCCCTGGCGCGCAGACCCTCCAGTCCCCGAGATGCCCGCAGCAACGC 299
| | | | |
Db 242 LysValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArg 261

QY 298 TTTGCCCACTTCCAGGAGCTGCCCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAG 357
| | | | |
Db 262 PheAlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLys 281

QY 358 CAAGTGCCTGGTTTCTGCAGCTGGCGCGGAGGACACAGATCGCCCTCTCGAAGCATCC 417
| | | | |
Db 282 GlnValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSer 301

QY 418 ACTATCGAGATCATCTGCTAGAGACGCGAGCTACACACAGAGACAGAGTGTATC 477
| | | | |
Db 302 ThrIleGluIleMetLeuLeuGluThrAlaArgArgTyrAenHisGluThrGluCysIle 321

QY 478 ACCTTCTTCAG-GACTTTCACCTACAGCAGGACGACTTCCACCGTCAGGCTGCAGGTG 536
| | | | |
Db 322 ThrPheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnVal 341

QY 537 GAGTTCATCAACCCCATCTTCGAGTTCTCGCGGGCCATCGCGCGCTGGGCTGCACAC 596
| | | | |
Db 342 GluPheIleAenProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAsp 361

QY 597 GCTGAGTAGCCCTGCTCATCCCATCAACATCTTTCGGCGGACCGCCCAACCTGCAG 656
| | | | |
Db 362 AlaGluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProasnValGln 381

QY 657 GAGCGCGCGCGCTGGAGCGTTTCAGCAGCGCCCTACGTGGAGCGCTGTGTCTACACG 716
| | | | |
Db 382 GluProSerArgValGluAlaLeuGlnProTyrValAspAlaLeuLeuSerTyrThr 401

QY 717 CGCATCAAGAGCCCGCAGCAGCAGCTGGCG 746
| | | | |
Db 402 ArgIleLysArgProGlnAspGlnLeuArg 411

RESULT 7

OSBIN1_BOVIN PRELIMINARY; PRT; 455 AA.

ID OSBIN1_BOVIN PRELIMINARY; PRT; 455 AA.
AC OSBIN1;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DE Nuclear receptor subfamily 1, group H, member 2.
GN Name=NR1H2;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RX NUCLEOTIDE SEQUENCE.
RP TISSUE=Pooled;
RC MEDLINE=21180013; PubMed=11282978; DOI=10.1101/gr.170101;
RA Smith T.P.L., Grosse W.M., Freking B.A., Roberts A.J., Stone R.T.,

CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.

DR EMBL; BT021903; AAX46750.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:steroid hormone binding; IEA.
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 498 AA; 54139 MW; 29CED3E4E6F4156 CRC64;

Alignment Scores:
Pred. No.: 7,09e-74 Length: 498
Score: 1144.50 Matches: 235
Percent Similarity: 94.80% Conservative: 2
Best Local Similarity: 94.00% Mismatches: 10
Query Match: 83.30% Indels: 4
DB: Gaps: 2

US-10-712-629B-2 (1-746) x Q58CP4_BOVIN (1-498)

QY 1 AGATTTCGGAACAGCAGCAGCAGGAGTCACTACAGTCGAGTCGAGTCGAGTCGAGCGG 60
DB 164 LysileArglyseGlnGlnGln-----GlnGlnGlnGlnSerSerProThrGlyPro 181
QY 61 CAGGCGCAGCAGCAGCTCA---GCCTCTGGGCGCTGGGGCTTCCCTCGTGATCTGAGGCA 117
DB 182 GlyValSerSerSerProAlaSerGlyProGlyAlaSerProGlyGlySerAspGly 201
QY 118 GCGCAGCAGGCTCCGGGAGCGAGGGTGTCCAGCTAACACGGCTCAAGAACTAATG 177
DB 202 GlyGlyGlnGlySerGlyGlyGlyGlyValGlnLeuThrAlaGlnGlnLeuMet 221
QY 178 ATCAGCAGTGTGTGGGGCCCACTGCAGTCAACAAAGCTCTCTCCGACGAGCCC 237
DB 222 IleGlnGlnLeuValAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnPro 241
QY 238 AAGATCAGCGCTGGCGCTGGCGGAGAGACCCAGTCGCCAGATGCCGCCAGCAACGC 297
DB 242 LysValThrProThrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArg 261
QY 298 TTGCCCATTACGAGCTGGCGCATCTCTAGTCCAGAGATCGTGGACTTCGCTAAG 357
DB 262 PheAlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLys 281
QY 358 CAAGTGCCTGTTCTCTGAGCTGGCGGAGGACAGATCGCCCTCTGAGGAGCATCC 417
DB 282 GlnValProGlyPheLeuGlnLeuGlyArgGlyAspGlnIleAlaLeuLeuLysAlaSer 301
QY 418 ACTATCGAGATCATGCTGTAGAGACGCGCGCTTACACCAACGAGACAGAGTGTATC 477
DB 302 ThrIleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIle 321
QY 478 ACCTTCTTAG-GACTTCACCTACAGCAAGAGCAGACTTCCACCGTCGAGCGCTCGAGGTG 536
DB 322 ThrPheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnVal 341
QY 537 GAGTTCATCAACCCATCTTCGAGTCTCGGGCCATCGGGCTGGCGCTGGAGCAG 596
DB 342 GluPheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAsp 361
QY 597 GCTGAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGGCGCAGCCGCCCAACGTCGAG 656
DB 362 AlaGluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGln 381
QY 657 GAGCCGGCGCGTGGAGGGCTTCACAGACGCCCTACCTGAGGGCGCTGCTGCTACAG 716
DB 382 GluProSerArgValGluAlaLeuGlnGlnProTyrValAspAlaLeuLeuSerTyrThr 401
QY 717 CGCATCAAGAGCGCGAGGACCACTCGCG 746
DB 402 ArgIleLysArgProGlnAspGlnLeuArg 411

RESULT 9

NR1H2 RAT STANDARD; . PRT; 446 AA.
AC Q62755; Q62694;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Oxyesterols receptor LXR-beta (liver X receptor beta) (Nuclear orphan
DE receptor LXR-beta) (Ubiquitously-expressed nuclear receptor) (UR)
DE (Orphan nuclear receptor OR-1).
DE Name=Nr1h2; Synonyms=Lxrb;
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95199298; PubMed=7892230;
RA Teboul M., Enmark E., Li Q., Wikstrom A.C., Pelto-Huikko M.,
Gustafsson J.-A.;
RT "OR-1, a member of the nuclear receptor superfamily that interacts
RT with the 9-cis-retinoic acid receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:2096-2100(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Vagina;
RX MEDLINE=95062154; PubMed=7971966;
RA Song C., Kokontis J.M., Hiipakka R.A., Liao S.;
RT "Ubiquitous receptor: a receptor that modulates gene activation by
RT retinoic acid and thyroid hormone receptors.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:10809-10813(1994).
CC -!- FUNCTION: Orphan receptor. Binds preferentially to double-stranded
CC oligonucleotide direct repeats having the consensus half-site
CC sequence 5'-AGTCA-3' and 4-nt spacing (DR-4).
CC -!- SUBUNIT: Forms a heterodimer with RXR.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR1
CC subfamily.
CC -!- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
DR EMBL; U20389; AAA69522.1; -; mRNA.
DR EMBL; U14533; AAA52361.1; -; mRNA.
DR PIR; I59354; I59354.
DR HSP; P55065; LP8D.
DR SNR; O62755; 205-445.
DR TRANSFAC; T04451; ...
DR Ensembl; ENSRNOG0000019812; Rattus norvegicus.
DR RGD; 61906; Nr1h2.
DR InterPro; IPR001628; Hrmn_rcpt_DNA_bd.
DR InterPro; IPR000536; Hrmn_rcpt_lig_bd.
DR InterPro; IPR001723; Stdhrn_rcptor.
DR InterPro; IPR000324; Vitd_rcptor.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf_C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
DR PROSITE; PS01030; NUCLEAR_REC_DBD_2; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
Transcription regulation; Zinc; Zinc-finger.

```
FT DNA_BIND 75 152 Nuclear receptor.
FT ZN_FING 78 98 NR C4-type.
FT ZN_FING 116 140 NR C4-type.
FT REGION 216 446 Ligand-binding (Potential).
FT COMPTAS 162 168 Poly-Gln.
FT COMPTAS 169 172 Poly-Pro.
FT CONFLICT 33 33 G -> V (in Ref. 2).
FT CONFLICT 52 54 Missing (in Ref. 2).
FT CONFLICT 219 219 A -> V (in Ref. 2).
SQ SEQUENCE 446 AA; 49736 MW; 13DF6DC2F0F5FA4D CRC64;

Alignment Scores:
Pred. No.: 9.5e-71
Score: 1101.00
Percent Similarity: 92.77%
Best Local Similarity: 90.36%
Query Match: 80.13%
DB: 1
Gaps: 1

US-10-712-629B-2 (1-746) x NR1H2_RAT (1-446)
QY 1 AAGATTCCGAAACAGCAGCAGCAGGAGTACAGTCACAGTCGCGAGTCACCTGTGGGGCCG 60
Db |||||:|||||
160 LysileinylsGlnGlnGln-----GlnProProProProThr 173
QY 61 CAGGCGCAGCAGCAGCAGCCTCTGGCGCTTCCCTGGTGATCTGAGGAGGC 120
Db :|||:|||||
174 GluProAlaSerGlySerAlaArgProAlaAlaSerProGlyThrSerGluAlaSer 193
QY 121 AGCAGGCTCCGGGAGCGAGGCTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db |||||:|||||
194 SerGlnGlySerGlyGluGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 213
QY 181 CAGCAGTTGTCGCGCCCACTGCAGTCGACCAACAGCGCTCTCCGACGAGCCCAA 240
Db |||||:|||||
214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTCACGCTTGGCGCTTGGCGCAGACCCCGAGTCCCGAGATCCCGCCAGCAGCAACGCTT 300
Db |||||:|||||
234 ValThrProThrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 253
QY 301 GCCACTTCAGGAGTGGCATCATCTCAGTCCAGGAGATCTGACTGCTGCTAAGCAA 360
Db |||||:|||||
254 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
QY 361 GTGCTGTTCTCTGAGTGGCGGAGGAGCAGATCGCTCTCTGAGGAGCATCACT 420
Db |||||:|||||
274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY 421 ATCGAGATCATGCTGATAGACAGCAGCGCTACAAACACGAGACAGAGTGTATCACC 480
Db |||||:|||||
294 IleGluIleMetLeuLeuGluThrAlaArgTyrAsnHisGluThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTCACCTACAGCAGGAGCAGCTTCCACCGTGCAGCGCTGCGAGTGGAG 539
Db |||||:|||||
314 PheLeuLysAspPheThrThrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGCATCGCGGCTGCGCTGCGAGCAGCT 599
Db |||||:|||||
334 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAla 353
QY 600 GAGTAGCGCTGTCTCATCGCATCAACATCTTCTCGCGCGACCGCGCCCAACGTCGAGGAG 659
Db |||||:|||||
354 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY 660 CCGGGCGGTGAGCGGTGAGCAGCCCTACGTGAGGCGCTGTCTGTCTACACGCC 719
Db |||||:|||||
374 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerThrArg 393
QY 720 ATCAGAGCGCGCAGCAGCAGCTGCCG 746
Db |||||:|||||
394 IleLysArgProGlnAspGlnLeuArg 402
```

```
RESULT 10
NR1H2_MOUSE
ID NR1H2_MOUSE STANDARD; PRT; 446 AA.
AC Q60644;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Oxyterols receptor LXR-beta (Liver X receptor beta) (Nuclear orphan
DE receptor LXR-beta) (Ubiquitously-expressed nuclear receptor) (Retinoid
DE X receptor interacting protein No.15).
GN Names=Nr1h2; Synonyms=Lxrb, Rlp15, Unr, Unr2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=95280959; PubMed=7760852; DOI=10.1210/me.9.1.72;
RA Seol W., Choi H.S., Moore D.D.;
RT "Isolation of proteins that interact specifically with the retinoid X
RT receptor: two novel orphan receptors.";
RL Mol. Endocrinol. 9:72-85(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=129/SvJ; TISSUE=Liver;
RX MEDLINE=20156373; PubMed=10675617; DOI=10.1016/S0378-1119(99)00555-7;
RA Alberti S., Steffensen K.R., Gustafsson J.-A.;
RT "Structural characterisation of the mouse nuclear oxysterol receptor
RT genes LXRalpha and LXRbeta.";
RL Gene 243:93-103(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Orphan receptor. Binds preferentially to double-stranded
CC oligonucleotide direct repeats having the consensus half-site
CC sequence 5'-AGGTCA-3' and 4-nt spacing (DR-4).
CC -!- SUBUNIT: Forms a heterodimer with RXR.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR1
CC subfamily.
CC -!- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC EMBL; U09419; AAC52164.1; -, mRNA.
```


DR EMBL; AJ132602; CAB51924.1; -; Genomic_DNA.
 DR EMBL; BC066025; AAH66025.1; -; mRNA.
 DR PIR; I49021; I49021.
 DR HSSP; P55055; 1P8D.
 DR SMR; Q60644; 205-445.
 DR TRANSFAC; T04467; -.
 DR Ensembl; ENSMUSG0000060601; Mus musculus.
 DR MGI; MGI:1352463; Nr1h2.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IDA.
 DR GO; GO:0003700; F:transcription factor activity; IDA.
 DR GO; GO:004255; P:cellular lipid metabolism; IDA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IDA.
 DR InterPro; IPR001628; Hrmn rcpt DNA bd.
 DR InterPro; IPR000536; Hrmn rcpt lig bd.
 DR InterPro; IPR001723; Strdhmrn receptor.
 DR InterPro; IPR000324; VitD_receptor.
 DR Pfam; PF00104; Hormone_recep; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STRDHORMNER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR PRINTS; PR00350; VITAMINDR.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOL1; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
 DR PROSITE; PS1030; NUCLEAR_REC_DBD_2; 1.
 DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger;
 FT DNA_BIND 75 152 Nuclear receptor.
 FT ZN_FING 78 98 NR C4-type.
 FT ZN_FING 116 140 NR C4-type.
 FT REGION 216 446 Ligand-binding (Potential).
 FT COMPTIAS 162 168 Poly-Gln.
 FT COMPTIAS 169 172 Poly-Pro.
 SQ SEQUENCE 446 AA; 49720 MW; 73153B635302C9DF CRC64;

 Alignment Scores:
 Pred. No.: 1.44e-70 Length: 446
 Score: 1098.50 Matches: 224
 Percent Similarity: 93.57% Conservative: 9
 Best Local Similarity: 89.96% Indels: 15
 Query Match: 79.95% Gaps: 2
 DB: 1

US-10-712-629b-2 (1-746) x NR1H2_MOUSE (1-446)
 QY 1 AAGATTCGGAACAGCAGCAGCAGGAGTCACAGTCACAGTCGCGAGTCACCTGTGGGCGG 60
 DB 155 GlnIleArgLysLysArgIleGlnLys---GlnGlnGlnGlnGlnProProProSer 173
 QY 61 CAGGGCAGCAGCAGCAGCAGCTCTGGGCGCTTCCCGCTGGGCTTCCCGTGGATCTGAGGCAGGC 120
 DB 174 GluProAlaIaIaSerSerSerGlyArgProAlaAlaSerProGlyThrSerGluAlaSer 193
 QY 121 AGCCAGGGCTCCGGGGAGCGAGGGTGTCCAGCTAACAGCGGCTCAAGAACTAATGATC 180
 DB 194 SerGlnGlySerGlyGluGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 213
 QY 181 CAGCAGTGTGGGGCGGCGCACTGCAGTCGCAACAAAGCTCTCTCCGACCGCCCAAA 240
 DB 214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
 QY 241 GTACAGCCCTGGGCGGCGCAGACCCCGAGATGCCCGGATGCCCGCCAGCAACGCTTT 300
 DB 234 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 253
 QY 301 GCCCACTTCAGGAGCTGGGCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTTAACAA 360
 DB 254 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
 QY 361 GTGCTGTGTTCTGCAGCTGGGCGGAGGAGCAGATCGCCCTCCCTGAGGCGATCCACT 420

Db 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
 QY 421 ATCGAGATCATGCTCTAGACACGCCAGGCGCTTCAACCCAGACAGAGTGTTATCACC 480
 Db 294 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
 QY 481 TTCCTTGAG-GACTTCACCTACAGCAGCAGCAGCTTCCACCGTCGAGGCGCTGCGAGTGGAG 539
 Db 314 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
 QY 540 TTCATCAACCCCATCTTCAGTTCTCGGGGGCCATCGCGCGCTGGCGCTGGACGACGCT 599
 Db 334 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla 353
 QY 600 GAGTACGCGCTCTCATCGCCATCAACATCTTCTCGCGCGACCGCGCCACGCTGCAGGAG 659
 Db 354 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
 QY 660 CCGGGCGGCTGGAGCGCTTGCAGCAGCCCTACGTGAGGCGCTGCTGTCTACACGCGC 719
 Db 374 ProSerArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 393
 QY 720 ATCAAGAGCGCGCAGCAGCAGCTGCGC 746
 Db 394 IleLysArgProGlnAspGlnLeuArg 402
 RESULT 11
 Q8BP65_MOUSE PRELIMINARY; PRT; 443 AA.
 AC Q8BP65_MOUSE 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched
 DE library, clone:5730478C17 product:nuclear receptor subfamily 1, group
 DE H, member 2, full insert sequence.
 DE Name=Nr1h2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehii Y.,
 RA Arakawa T., Harai A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant M.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bareh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Kamiya M., Lee N.H.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).

[3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN (4)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN (5)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN (6)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saio K., Saitoh K., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Tova T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; AK077620; BAC36906.1; -; mRNA.
DR HSSP; P55055; LPQ6.
DR SMR; Q89P65; 202-442.
DR MGI; MGI:1352463; Nr1h2.
DR GO; GO:0008634; C:nucleus; IDA.
DR GO; GO:0004879; Filigand-dependent nuclear receptor activity; IDA.
DR GO; GO:0003700; P:transcription factor activity; IDA.
DR GO; GO:0044255; P:cellular lipid metabolism; IDA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IDA.
DR InterPro; IPR000536; Hrmn_recept_lig.
DR InterPro; IPR001723; Stdhrmn_receptor.
DR InterPro; IPR000324; VitD_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STDRHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 443 AA; 49394 MW; 362DCB2D21034F29 CRC64;

Alignment Scores:
Pred. No.: 2,36e-70 Length: 443
Score: 1095.50 Matches: 224
Percent Similarity: 93.57% Conservative: 9
Best Local Similarity: 89.96% Mismatches: 15
Query Match: 79.73% Indels: 2
DB: 2 Gaps: 1
US-10-712-629B-2 (1-746) x Q8BP65_MOUSE (1-443)
QY 1 AGATTTCGGAACAGCAGCAGCAGGAGTACACATCAGTCGCGAGTCACCTGTGGGCGC 60
Db 152 GlnileArgLysLysArgileGlnLys---GlnGlnGlnGlnProProProSer 170
QY 61 CAGGCGCAGCAGCAGCAGCTCAGCCTCTGGGCGCTTCCCTCGTGGAGTCTGAGGCGAGC 120
Db 171 GluProAlaAlaSerSerGlyArgProAlaAlaSerProGlyThrSerGluAlaSer 190
QY 121 AGCAGGCGCTCCGGGGAAGCGGAGGCTGTCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db 191 SerGlnGlySerGlyGluGlyGluGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 210
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAAGTGCACAAACAGCTCTCTCCGACCCAGCCCAA 240
Db 211 GlnGlnLeuValAlaAlaGlnLeuGlnCysAenLysArgSerPheSerAepGlnProLys 230
QY 241 GTCACGCCCTGGCGCCCTGGCGGCGAGACCCCGCAGTCCCGAGATCCCGCCAGCAACGCTTT 300
Db 231 ValThrProTrpProLeuGlyAlaAepProGlnSerArgAepAlaArgGlnGlnAtgPhe 250
QY 301 GCCCACTTCACGAGCTGCCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAGCA 360
Db 251 AlaHisPheThrGluLeuAlaIleIleSerValGlnGlnIleValAspPheAlaLysGln 270
QY 361 GTGCTTGGTTTCTCAGCTGGCGCGGAGGAGCAGATCGCCCTCTCTGAAGGATCCACT 420
Db 271 ValProGlyPheLeuGlnLeuGlyArgGluAepGlnIleAlaLeuLeuLysAlaSerThr 290
QY 421 ATCGAGATCATGCTGTAGACAGCAGCGCGCTACACCCAGCAGAGAGTGTATCACC 480
Db 291 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 310
QY 481 TTCTTGAG-GACTTTCACCTACAGCAGGAGCAGCTCCACCGTCAGCGCTCGAGGTGGAG 539
Db 311 PheLeuLysAspPheThrTyrSerLysAspPhePheHisArgAlaGlyLeuGlnValGlu 330
QY 540 TTCATCAACCCCATCTTCAGTTCTCGCGGCGCATCGCGCGCTGGGCGCTGGACGCGCT 599
Db 331 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAepAepAla 350
QY 600 GAGTACGCCCTCTCATCTCATCCCATCAACATCTTCTCGCGCGAGCCCGCCCAACGTGCGAGG 659
Db 351 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAepArgProAsnValGlnGlu 370
QY 660 CCGGCGCGGTGGAGCGGTTCAGCAGCAGCGCTACGTGGAGCGCGCTGTGCTGTCTACACGCGC 719
Db 371 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 390
QY 720 ATCAAGAGCGCGCAGGACCGAGTGGCG 746
Db 391 IleLysArgProGlnAepGlnLeuArg 399
RESULT 12
QSREL4_PONPY
ID QSREL4_PONPY PRELIMINARY; PRT; 351 AA.
AC QSREL4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp468A0622.
GN Name=DKFZp468A0622;
OS Pongo pygmaeus (Orangutan).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Heart;
 RG The German cDNA Consortium;
 RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaipp A.,
 RA Newes H.W., Weil B., Amid C., Oeanger A., Fobo G., Han M., Wiemann S.;
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
 DR EMBL; CR857510; CAH89793.1; -; mRNA.
 DR SMR; Q5REL4; 140-350.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0003708; F:steroid hormone receptor activity; IEA.
 DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR InterPro; IPR000536; Hrmn_recept_lig.
 DR InterPro; IPR003078; Rtnoid_receptor.
 DR InterPro; IPR001723; Stdhmn_receptor.
 DR InterPro; IPR000324; Vitd_receptor.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00104; Hormone_recep; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR01292; RETNOICACIDR.
 DR PRINTS; PR00398; STRDHORMONR.
 DR PRINTS; PR00047; STROIDFINGER.
 DR PRINTS; PR00350; VITAMINDR.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOL1; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
 KW DNA-binding; Hypothetical protein; Nuclear protein;
 KW Receptor; Transcription; Transcription regulation; Zinc; zinc-finger.
 SQ SEQUENCE 351 AA; 39475 MW; 175C907978A0247C CRC64;
 Alignment Scores:
 Pred. No.: 1.68e-65 Length: 351
 Score: 1028.00 Matches: 214
 Percent Similarity: 86.35% Conservative: 1
 Best Local Similarity: 85.94% Mismatches: 4
 Query Match: 74.82% Indels: 31
 DB: 2 Gaps: 1
 US-10-712-629b-2 (1-746) x Q5REL4_PONPY (1-351)
 QY 1 AAGATTGGAAACAGCAGCAGCAGGAGTTCACAGTCGAGTCACCTGTGGGGCGG 60
 DB 89 LysileargylsGlnGlnGlnGlnSerGlnSerGlnSerGlnSerProAlaGlyPro 108
 QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCCCTGGGGCTTCCCTCGTGGAGCTGAGCGAGC 120
 DB 109 GlnGlySerCysSerSerThrSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 128
 QY 121 AGCAGGGCTCCGGGAGCGAGGGTGTCCAGTACACGGCTCAAGACTAATGATC 180
 DB 129 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetile 148
 QY 181 CAGCAGTTGGTGGGGCCCAACTGCAGTCAACAAAGCTCTTCTCCGACCGAGCCCAAA 240
 DB 149 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProlys 168
 QY 241 GTACAGCCCTGGCCCTTGGGGCGCAGACCCCGAGTCCCGAGATCCCGCAGCAACGCTTT 300
 DB 169 ValThr----- 170
 QY 301 GCCCACTTCAGGAGCTGGCCATCATCTCAGTCAGGAGATCGGGACTTCGCTTAAGCAA 360
 |||||

Db 171 -----GluIleValaspPheAlalysGln 178
 QY 361 GTGCCTGGTTTCTCAGCTGGCGGAGGAGCAGATCGCCTCTCTGAAGCATCCACT 420
 |||||
 Db 179 ValProGlyPheLeuGlnLeuGlyArgGluAaspGlnIleAlaLeuLeuLysAlaSerThr 198
 QY 421 ATCGAGATCATCTCTAGACACCCAGCGCTACAACCAGCAGACAGAGTGATCACC 480
 |||||
 Db 199 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 218
 QY 481 TTCTTTGAG-GACTTCACCTACAGCAGCAGCTTCCACCGTCGAGGCTCGCAGGTGGAG 539
 |||||
 Db 219 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 238
 QY 540 TTTCATCAACCCCATCTTCGAGTTCTCGCGGGCCATCGCGGCTGGCGCTGAGCAGCT 599
 |||||
 Db 239 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAaspAla 258
 QY 600 GAGTACGGCTCTCATGCCATCAACATCTTCTCGCGCGAGCGGCCCAACGTGAGGAG 659
 |||||
 Db 259 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAaspArgProAsnValGlnGlu 278
 QY 660 CCGGGCGCGTGGAGCGTTGCGAGCGCTACGTGGAGCGCTGTCTCTACACGCGC 719
 |||||
 Db 279 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 298
 QY 720 ATCAAGAGCGCGCAGGACCGCTGCGC 746
 |||||
 Db 299 IleLysArgProGlnAaspGlnLeuArg 307
 RESULT 13
 Q6GMA2 XENLA
 ID Q6GMA2 XENLA PRELIMINARY; PRT; 441 AA.
 AC Q6GMA2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE MGC81974 protein.
 GN Names=MGC81974;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins B., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Strapletenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Butterfield Y.S.N., Krzywinaki M.I., Skaleka U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Kidney;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

Richardson P.;
 "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 initiative";
 Dev. Dyn. 225:384-391(2002).
 [3].
 RN NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RA Klein S., Gerhard D.S.;
 Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
 DR EMBL; BC074169; A474169.1; -, mRNA.
 DR SMR; Q6GMA2; 200-440.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004884; F:ecdysteroid hormone receptor activity; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0005496; F:steroid binding; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR InterPro; IPR003069; Ecdystd_receptor.
 DR InterPro; IPR000536; Hormon_recept_lig.
 DR InterPro; IPR001723; Sterhmn_receptor.
 DR InterPro; IPR003324; VitD_receptor.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00104; Hormone_recep; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR01283; ECDYSTEROIDR.
 DR PRINTS; PR00398; STRDHORMONER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR PRINTS; PR00350; VITAMINDR.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOLI; 1.
 DR SMART; SM00399; ZNF_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
 DR PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
 DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger.
 SQ SEQUENCE 441 AA; 50229 MW; AF7FD8B8F11741CB CRC64;

Alignment Scores:
 Pred. No.: 9e-54 Length: 441
 Score: 865.00 Matches: 184
 Percent Similarity: 77.82% Conservatives: 9
 Best Local Similarity: 74.19% Mismatches: 35
 Query Match: 62.95% Indels: 21
 DB: 2 Gaps: 1

US-10-712-629B-2 (1-746) x Q6GMA2_XENLA (1-441)

QY 1 AAGATTTCGGAACAGCAGCAGCAGCAGTCCACAGTCACAGTCACAGTCGAGTCACCTGTGGGCGC 60
 Db 169 LysileArgLysGlnAspGluaspValThrArgSerSerAlaLeuValProPro 188
 QY 61 CAGGCGCAGCAGCAGTCACGCTTCGCGCTGGGCTTCCCTGGTGATCTGAGCAGCAGC 120
 Db 189 SerProCysMetLeuSerGln----- 195
 QY 121 ACCCAGGCTCCGGGAGCGCGAGGTGTCAGCTAACAGCGCTCAAGAACTAATGATC 180
 Db 196 -----GluValValGlnLeuThrProGlnGlnGlnLysMetile 208
 QY 181 CAGCAGTGTGGCGGCCCAACTGAGTGCACAAACGCTCTCTCCGACAGCAGCCCAA 240
 Db 209 GluGlnLeuValSerAlaGlnGlnGlnCysAsnLysArgSerPheSerAspGlnProlys 228
 QY 241 GTCAGCCTGCGCCCTGGCGCAGACCCCGAGTCCCGAGATGCCCGCAGCAACGCTTT 300
 Db 229 ValThrProTriProGlyThrAspProAsnSerArgGluAlaArgGlnGlnArgPhe 248
 QY 301 GCCCACTTCACGAGCTGCCCATCATCTCAGTCCAGGAGATCGTGGACTTCCTAAGCAA 360
 Db 249 AlaHisPheThrGluLeuAlailelleSerValGlnGlnValAspPheAlaLysGln 268

QY 361 GTGCTGTGGTTCTTCGAGCTGGCGGAGGAGCAGATCGCCCTCTCTGAGGATCCACT 420
 Db 269 ValProGlyPheLeuGluLeuSerArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 288
 QY 421 ATCGAGATCATGCTGTAGACAGCCAGCGCTACACACGAGACAGAGTGTATCACC 480
 Db 289 IleGluIleMetLeuLeuGluThrAlaArgArgTyAsnHisGluThrGluCysIleThr 308
 QY 481 TTCTTGAG-GACTTCACCTACAGCAGGAGCAGCTTCCACCGTCGAGCGCTGCAAGTGGAG 539
 Db 309 PheLeuLysAspPheThrTySerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 328
 QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGGCCATCGCGCGCTCGGACGACGCT 599
 Db 329 PheIleAsnProIlePheGluPheSerArgGlyMetArgGlnMetGlnLeuAspAspAla 348
 QY 600 GAGTACGCGCTGTCTATCCCATCAACATCTTCTCGCGCGACCGCGCCCAACGTCGAGGAG 659
 Db 349 GluTyAlaLeuLeuIleAlaileAsnIlePheSerAlaAspArgProAsnValLeuAsn 368
 QY 660 CCGGCGCGGTGAGCGGTTCGAGCAGCAGCCCTACGTGGAGCGCTGTCTCTACACGCGC 719
 Db 369 HisGlnHisValGluAsnLeuGlnLeuProTyValGluAlaLeuHisSerTyThrArg 388
 QY 720 ATCAAGAGCGCGCAGGACCGAGCTG 743
 Db 389 IleLysArgProGlnAspHisLeu 396

RESULT 14
 Q56A56 BRARE PRELIMINARY; PRT; 412 AA.
 AC Q56A56;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Hypothetical protein zgc:113339.
 GN Name=zgc:113339;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Singapore local strain; TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Haigh F.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haigh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Hellon E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smaluk D.E.,
 RA Schnerch A., Schain J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Singapore local strain; TISSUE=Embryo;
 RL NIH MGC Project;
 RG Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.

RA Alberti S., Steffensen K.R., Gustafsson J.-A.;

genes LXRA and LXRBeta." ;
Gene 243:93-103(2000).
[2]
NUCLEOTIDE SEQUENCE.
Chen Y.E., Horiuchi M., Dzau V.J.;
Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: Orphan receptor. Interaction with RXR shifts RXR from
its role as a silent DNA-binding partner to an active ligand-
binding subunit in mediating retinoid responses through target
genes defined by LXRES. LXRES are DR4-type response elements
characterized by direct repeats of two similar hexanucleotide half-
sites spaced by four nucleotides. Plays an important role in the
regulation of cholesterol homeostasis.
-1- SUBUNIT: Heterodimer of LXRA and RXR.
-1- SUBCELLULAR LOCATION: Nuclear (Potential).
-1- SIMILARITY: Belongs to the nuclear hormone receptor family. NRI
subfamily.
-1- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; AJ132599; CAB51952.1; -; Genomic_DNA.
EMBL; AJ132600; CAB51952.1; JOINED; Genomic_DNA.
EMBL; AJ132601; CAB51923.1; -; mRNA.
EMBL; AF085745; AAD16050.1; -; mRNA.
HSSP; P55055; 1P8D.
SMR; Q9Z0Y9; 204-444.
TRANSFAC; T04430; -.
Ensembl; ENSMUSG00000002108; Mus musculus.
MGI; MGI:1352462; Nr1h3.
GO; GO:0005634; C:nucleus; IDA.
GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IDA.
GO; GO:0003700; F:transcription factor activity; IDA.
GO; GO:0004255; P:cellular lipid metabolism; IDA.
GO; GO:0016481; P:negative regulation of transcription; IMP.
GO; GO:0006355; P:regulation of transcription, DNA-dependent; IDA.
InterPro; IPR003069; Ecdystrd receptor.
InterPro; IPR001628; Hrmn_rcbt_DNA_bd.
InterPro; IPR000536; Hrmn_rcbt_lig_bd.
InterPro; IPR001723; Stdhrn_receptor.
Pfam; PF00104; Hormone_recep; 1.
Pfam; PF00105; zf-C4; 1.
PRINTS; PR01283; ECDYSTEROIDR.
PRINTS; PR00398; STRDHORMONER.
PRINTS; PR00047; STROIDFINGER.
ProDom; PD000035; Znfc4steroid; 1.
PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.

DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW

KW Transcription regulation; Zinc; Zinc-finger.
FT DNA BIND 93 168 Nuclear receptor.
FT ZN_FING 96 116 NR C4-type.
FT ZN_FING 132 156 NR C4-type.
FT REGION 213 432 Ligand-binding (Potential).
FT CONFLICT 399 R -> P (in Ref. 2).
SQ SEQUENCE 445 AA; 50477 MW; CC9A8DF38D935593 CRC64;

Alignment Scores:

Pred. No.:	1-89e-43	Length:	445
Score:	721.50	Matches:	152
Percent Similarity:	75.40%	Conservative:	35
Best Local Similarity:	61.29%	Mismatches:	38
Query Match:	52.51%	Indels:	24
DB:	1	Gaps:	2

US-10-712-629B-2 (1-746) x NR1H3_MOUSE (1-445)

```
QY 1 AGATTTCGGAACAGCAGCAGCAGGAGTCACAGTCACAGTCGCGAGTCACCTGTGGGCGCG 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 176 LysLeuLysArgGlnGluGlnAlaGlnAlaThrSerValSerPro----- 192

QY 61 CAGGGCAGCAGCAGCAGCTCAGCCTCTGGGCCCTGGGGCTTCCCTCGTGATCTGAGGCAGGC 120
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 193 -----ArgValSerSerProGlnValSerPro----- 202

QY 121 AGCCAGGGCTCCGGGAAGCGAGGGTGTCCAGCTAACAGCGGCTCAAGAACTAATGATC 180
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 203 -----GlnLeuSerProGlnLeuGlyMetIle 212

QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAAGTCAGTGCACAAACGGTCTCTTCCGACCAGCCCAAA 240
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 213 GluLysLeuValAlaAlaGlnGlnCysAsnArgSerPheSerAspArgLeuArg 232

QY 241 GTCACGCCCTGGGCCCTGGCGGCGAGACCCCGTCCGAGATGCCCCGAGCAACCGCTTT 300
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 233 ValThrProTrpProIleAlaProAspProGlnSerArgGluAlaArgGlnArgPhe 252

QY 301 GCCCACTTCACGAGCTGCCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 253 AlaHisPheThrGluLeuAlaIleValSerValGlnGluIleValAspPheAlaLysGln 272

QY 361 GTGCTGTGTTCTCTGAGCTGGCGGAGGAGCACAGATCGCCCTCTGAAGGCATCCACT 420
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 273 LeuProGlyPheLeuGlnLeuSerArgGluAspGlnIleAlaLeuLeuLysThrSerAla 292

QY 421 ATCGAGATCATGCTGTAGACAGCAGCGCGCTACAACACGACGACAGAGTGTATCACC 480
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 293 IleGluValMetLeuLeuGluThrSerArgTyrAsnProGlySerGluSerIleThr 312

QY 481 TTCTTCGAG-GACTTCACCTACAGCAAGGAGGACTTCCACCGTCGAGCGCTGCGAGTGGAG 539
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 313 PheLeuLysAspPheSerTyrAsnArgGluAspPheAlaLysAlaGlyLeuGlnValGlu 332

QY 540 TTCATCAACCCCATCTTCTGAGTTCTCGCGGCGCATCGCGGCGCTGGACACGCGCT 599
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 333 PheIleAsnProIlePheGluPheSerArgAlaMetAsnGlnLeuGlnLeuAsnAspAla 352

QY 600 GAGTACGCCCTGTCTATCGCATCAACATCTTCTCGGCGGACCGCGCCAAACGTGACAGGAG 659
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 353 GluPheAlaLeuLeuIleAlaIleSerIlePheSerAlaAspArgProAsnValGlnAsp 372

QY 660 CCGGGCGCGGTGGAGCGTTGAGCAGACCCCTAGCTGGAGCGCTGCTGCTACACGCGC 719
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 373 GlnLeuGlnValGluArgLeuGlnHisThrTyrValGluAlaLeuHisAlaTyrValSer 392

QY 720 ATCAGAGGCGCGCAGCAGCAGCTG 743
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 393 IleAsnHisProHisAspArgLeu 400
```

Search completed: November 25, 2005, 22:30:18
Job time : 231 secs

This Page Blank (Page 2)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2005, 22:17:28 ; Search time 37 Seconds
(without alignments)

3333.841 Million cell updates/sec

Title: US-10-712-629B-2

Perfect score: 1374

Sequence: 1 aagattcggaacagcagca.....ggccgcaggaccagtgcgc 746

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DRV=xlp
-Q/cgn2_1/USPTO_spool_p/US10712629/runat 25112005 143037 11014/app_query.fasta_1.903
-DB=issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10712629 @CNC 1.1 71 @runat 25112005 143037 11014 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5 COMB.pcp.*
2: /cgn2_6/prodata/1/iaa/6 COMB.pcp.*
3: /cgn2_6/prodata/1/iaa/H COMB.pcp.*
4: /cgn2_6/prodata/1/iaa/PCTUS COMB.pcp.*
5: /cgn2_6/prodata/1/iaa/RE COMB.pcp.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1231	89.6	461	1	US-08-330-518-2
2	1231	89.6	461	1	US-08-330-283-2
3	1231	89.6	461	1	US-08-646-248-2
4	1231	89.6	461	4	PCT-US95-13924-2
5	1231	89.6	461	4	PCT-US95-13931-2
6	1231	89.6	478	2	US-09-949-016-7717
7	1227	89.3	461	2	US-09-976-594-655
8	1223	89.0	460	1	US-08-342-411A-2
9	1215.5	88.5	460	2	US-10-329-668-4
10	1101	80.1	446	2	US-08-776-844-2
11	1101	80.1	446	2	US-09-909-325-2
12	1101	80.1	446	2	US-09-909-326-2

13	1095.5	79.7	446	1	US-08-372-652-3
14	1095.5	79.7	446	4	PCT-US95-16311-3
15	1094	79.6	443	1	US-08-342-411A-4
16	707.5	51.5	453	2	US-09-949-016-8058
17	707.5	51.5	453	2	US-09-949-016-8413
18	705.5	51.3	440	1	US-08-333-358-8
19	705.5	51.3	440	1	US-08-463-694-8
20	705.5	51.3	440	1	US-08-694-501-8
21	702.5	51.1	447	1	US-08-373-935-1
22	702.5	51.1	447	2	US-10-329-668-2
23	692	50.4	433	1	US-08-466-120-2
24	680	50.4	433	4	PCT-US94-07266-2
25	680	49.5	191	2	US-10-104-047-2537
26	323	23.5	496	2	US-10-087-167-86
27	321	23.4	478	2	US-10-087-167-82
28	321	23.4	541	2	US-10-087-167-12
29	298.5	21.7	510	2	US-10-087-167-80
30	298.5	21.7	599	2	US-10-087-167-127
31	296.5	21.6	521	2	US-10-087-167-90
32	294.5	21.4	536	2	US-10-087-167-14
33	294	21.4	507	2	US-10-087-167-78
34	293.5	21.4	516	2	US-10-087-167-88
35	293.5	21.4	556	2	US-09-564-418-11
36	292	21.3	516	2	US-10-087-167-72
37	291.5	21.2	279	2	US-10-087-167-8
38	291.5	21.2	475	2	US-10-087-167-129
39	291.5	21.2	499	2	US-10-087-167-70
40	291.5	21.2	505	2	US-10-087-167-94
41	291.5	21.2	553	2	US-10-087-167-105
42	291.5	21.2	556	2	US-10-087-167-2
43	291.5	21.2	588	2	US-10-087-167-123
44	291.5	21.2	599	2	US-10-087-167-148
45	291.5	21.2	602	2	US-10-087-167-143

ALIGNMENTS

RESULT 1
US-08-330-518-2
; Sequence 2, Application US/08330518
; Patent No. 5607967
; GENERAL INFORMATION:
; APPLICANT: Friedman, Eitan
; APPLICANT: Holloway, M. Katharine
; APPLICANT: Rodan, Gideon
; APPLICANT: Schmidt, Azriel
; APPLICANT: Vogel, Robert
; TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 East Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/330,518
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dolan, Catherine A.
; REGISTRATION NUMBER: 36,502
; REFERENCE/DOCKET NUMBER: 19316
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-4283
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:

```
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 461 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
US-08-330-518-2

Alignment Scores:
Pred. No.: 1,1e-104 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 1 Gaps: 0

US-10-712-629B-2 (1-746) x US-330-518-2 (1-461)
QY 1 AAGATTTCGGAACAGCAGCAGCAGGAGTCACTACAGTCGCGAGTCACCTGTGGGGCCG 60
Db 169 LysileArglysgInGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCTCAGCTCTGGGCTTGGGGCTTCCCTGGTGGATCTGAGGCAGGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCCGGGAAGCGGAGGTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetile 228
QY 181 CAGCAGTTGGTGGCGGCCAACTCGAGTCGACAAACAGCTCTCTCCGACAGCCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCCAGCCCTGGCCCTGGCGGAGGAGCCAGATCCGCTCTGAAAGGCATCCACT 300
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCCAGATCATGTCTAGACAGCCAGCGCTTACAAACCAGCAGACAGAGTGTATCACC 480
Db 309 IleGluileMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCTTGAG-CACVTCACCTACAGCAGGAGCAGCTTCCACCGTCGAGCGCTGCGAGTGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTCAATCAACCCCATCTTCGAGATTCCTCGCGGCCCATCGGGCTTGGGCTTGGAGCAGCT 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
QY 600 GAGTACGCCCTGTCTATCGCATCAACATCTTCTGGCCGACCGGCCCAACGTCGAGGAG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCCGCTGGAGCGTTCAGCAGCCCTTACGTGGAGCGCTGTCTCTACACCGCC 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuSerTyrThrArg 408
QY 720 ATCAAGAGGCGCAGGACCACTCGC 746
Db 409 IleLysArgProGlnAspGlnLeuArg 417
RESULT 2
```

```
US-08-330-283-2
/ Sequence 2, Application US/08330283
/ Patent No. 5679518
/ GENERAL INFORMATION:
/ APPLICANT: Friedman, Eitan
/ APPLICANT: Holloway, M. Katharine
/ APPLICANT: Rodan, Gideon
/ APPLICANT: Rutledge, Su Jane
/ APPLICANT: Schmidt, Azriel
/ APPLICANT: Vogel, Robert
/ TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
/ NUMBER OF SEQUENCES: 5
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Merck & Co., Inc.
/ STREET: 126 East Lincoln Avenue
/ CITY: Rahway
/ STATE: New Jersey
/ COUNTRY: US
/ ZIP: 07065-0907
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/330,283
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Dolan, Catherine A.
/ REGISTRATION NUMBER: 36,502
/ REFERENCE/DOCKET NUMBER: 19327
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (908) 594-4283
/ TELEFAX: (908) 594-4720
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 461 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
US-08-330-283-2

Alignment Scores:
Pred. No.: 1,1e-104 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 1 Gaps: 0

US-10-712-629B-2 (1-746) x US-330-283-2 (1-461)
QY 1 AAGATTTCGGAACAGCAGCAGCAGGAGTCACTACAGTCGCGAGTCACCTGTGGGGCCG 60
Db 169 LysileArglysgInGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCTCAGCTCTGGGCTTGGGGCTTCCCTGGTGGATCTGAGGCAGGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCCGGGAAGCGGAGGTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetile 228
QY 181 CAGCAGTTGGTGGCGGCCAACTCGAGTCGACAAACAGCTCTCTCCGACAGCCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCCAGCCCTGGCCCTGGCGGAGGAGCCAGATCCGCTCTGAAAGGCATCCACT 300
```


[illegible]

RESULT 3

```

US-08-646-248-2
; Sequence 2, Application US/08646248
; Patent No. 5939322
; GENERAL INFORMATION:
; APPLICANT: Friedman, Eitan
; APPLICANT: Holloway, M. Katharine
; APPLICANT: Rodan, Gideon
; APPLICANT: Rutledge, Su Jane
; APPLICANT: Schmidt, Azriel
; APPLICANT: Vogel, Robert
; TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 East Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,248
; FILING DATE: 14-MAY-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/330,283
; FILING DATE: 27-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Dolan, Catherine A.
; REGISTRATION NUMBER: 36,502
; REFERENCE/DOCKET NUMBER: 19327
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-4283

```

241 GTCAGCGCTGGCCCTGGCGGACAGCCCGAGTCCCGAGATGCCCGCAGCAACGCTTT 300
249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
301 GCCCACTTTCAGGAGCTGGCCATCATCTCAGTCAGGAGATCGTGGACTTCGCTAAGCAA 360
269 AlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLysGln 288
361 GTGCTGGTTTCTCGAGCTGGCGGAGGACAGATCGCCCTCTCTGAAGGCATCCACT 420
289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
421 ATCGAGATCATCTCTAGAGACAGCCAGGCGCTACAACACGAGACAGAGTATCACCC 480
309 IleGluIleMetLeuLeuGluThrAlaArgArgfyrAsnHisGluThrGluCysIleThr 328
481 TTCTTTGAG - GACTTCACCTTACAGCAAGGACGACTTCCACCGTCGAGGCTCGAGGTGGAG 539
329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
540 TTTCATCAACCCCATCTTCGAGTTCTCGGGGCGCATCGCGCGCTGGGCTGGACGCGCT 599
349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
600 GAGTAGCGCTCTCATGCCCATCAACATCTTCTCGCGGCGAGCCGCGCCCAACGTCAGGAG 659
369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
660 CCGGCGCGCTGGAGCGTTGCAGCAGCCCTACGTGGAGGCGCTGCTGCTCTACACGCGC 719
389 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
720 ATCAAGAGCGCCGAGACAGCTGCGC 746
409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 5
PCT-US95-13931-2
; Sequence 2, Application PC/TUS9513931
; GENERAL INFORMATION:
; APPLICANT: Friedman, Eitan
; APPLICANT: Holloway, M. Katharine
; APPLICANT: Rodan, Gideon
; APPLICANT: Schmidt, Azriel
; APPLICANT: Vogel, Robert
; TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 East Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13931
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Quagliato, Carol S.
; REGISTRATION NUMBER: 35,330
; REFERENCE/DOCKET NUMBER: 19316 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3809
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
PCT-US95-13924-2

Alignment Scores:
Pred. No.: 1,le-104 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 4 Gaps: 0

RESULT 4
PCT-US95-13924-2
; Sequence 2, Application PC/TUS9513924
; GENERAL INFORMATION:
; APPLICANT: Friedman, Eitan
; APPLICANT: Holloway, M. Katharine
; APPLICANT: Rodan, Gideon
; APPLICANT: Rutledge, Su Jane
; APPLICANT: Schmidt, Azriel
; APPLICANT: Vogel, Robert
; TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 East Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13924
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Quagliato, Carol S.
; REGISTRATION NUMBER: 35,330
; REFERENCE/DOCKET NUMBER: 19327 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3809
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
PCT-US95-13924-2

Alignment Scores:
Pred. No.: 1,le-104 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 4 Gaps: 0

US-10-712-629b-2 (1-746) x PCT-US95-13924-2 (1-461)

QY 1 AAGATTTCGAAACAGCAGCAGGAGTACAGTCAGTCGAGTCACCTGCGGCGCG 60
Db 169 LysIleArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGCGCAGCAGCTCAGCTCTGGGCGCTTGGGCTTCCCTGGTGGATCTGAGGCGGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGlyAlaGly 208
QY 121 AGCAGGCTCCGGGAGGAGGAGGTCTCAGCTAACAGGGCTCAAGACTAATGATC 180
Db 209 SerGlnGlySerGlyGlyGlyGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
QY 181 CAGCAGTTGGTGGGCGCCCAACTCAGTCAGTGCACAAACGCTCTTCTCCGACAGCCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248


```
US-10-712-629B-2 (1-746) x US-08-342-411A-2 (1-460)
QY 1 AAGATTTCGGAACAGCAGCAGCAGCAGTACAGTCCAGTCCAGTCCAGTCTGGGGCCG 60
DB 168 LysileargylsGlnGlnGlnGlnSerGlnSerGlnSerGlnSerProValGlyPro 187
QY 61 CAGGGCAGCAGCAGCAGTCCAGTCTGGGCCCTGGGGCTTCCCTGGTGGATCTGAGGCAGGC 120
DB 188 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 207
QY 121 AGCCAGGGCTCCGGGAGCGGGGTGTCCAGCTAACAGCGGCTCAAGAACTAATGATC 180
DB 208 SerGlnGlySerGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 227
QY 181 CAGCAGTTGCTGGCGGCCCACTGCAGTCAAAACAGCTCTCTCCGACCCAGCCCAAA 240
DB 228 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 247
QY 241 GTCACGCCCTGGGCCCTGGCGGAGACCCCACTGCAGTCCAGAGATCCCGCCAGCAACGCTTT 300
DB 248 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 267
QY 301 GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCTGGACTTCGCTAAGCAA 360
DB 268 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 287
QY 361 GTCCTGGTTTCTCAGCTGGCGGAGGACCCAGATCCGCTCTCGAAGGCATCCACT 420
DB 288 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 307
QY 421 ATCAGATCATGTCTAGAGACAGCGCGCTACAAACAGCGCTTCCACCCAGCAGAGAGTGTATCACC 480
DB 308 IleGluIleMetLeuLeuGluThrAlaArgTyrAsnHisGluThrGluCysIleThr 327
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGGACGACTTCCACCGTCCAGGCTGCGAGGTGGAG 539
DB 328 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 347
QY 540 TTCATCAACCCCATCTTCAGATTCTCGGGGCGCATCGGGGCTGGGGCTGGAGCAGCT 599
DB 348 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 367
QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGCGCCCAACGTCAGGAG 659
DB 368 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 387
QY 660 CCGGGCGGCTGAGGCTTGCAGCAGCCCTACGTGAGGCGCTGCTGCTTACACGCGC 719
DB 388 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 407
QY 720 ATCAAGAGCGCGCAGGACGAGTGGCG 746
DB 408 TyrLysArgProGlnAspGlnLeuArg 416
```

RESULT 9

```
US-10-329-668-4
; Sequence 4, Application US/10329668
; Patent No. 6696473
; GENERAL INFORMATION:
; APPLICANT: Martin Richard
; APPLICANT: Brenton Todd
; APPLICANT: Kahl Jeffrey Dean
; APPLICANT: Wang Tie-Lin
; TITLE OF INVENTION: HETEROCYCLIC MODULATORS OF NUCLEAR RECEPTORS
; FILE REFERENCE: 38205-3001
; CURRENT APPLICATION NUMBER: US/10/329, 668
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/342, 720
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 460
```

```
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-329-668-4
```

Alignment Scores:

```
Pred. No.: 2,9e-103 Length: 460
Score: 1215.50 Matches: 247
Percent Similarity: 99.20% Conservative: 0
Best Local Similarity: 99.20% Mismatches: 1
Query Match: 88.46% Indels: 2
DB: 2 Gaps: 1
```

US-10-712-629B-2 (1-746) x US-10-329-668-4 (1-460)

```
QY 1 AAGATTTCGGAACAGCAGCAGCAGGAGTACAGTCCAGTCCAGTCCAGTCTGGGGCCG 60
DB 169 Lysileargyls---GlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 187
QY 61 CAGGGCAGCAGCAGCAGCTCAGCTCTGGGCCCTGGGGCTTCCCTGGTGGATCTGAGGCAGGC 120
DB 188 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 207
QY 121 AGCCAGGGCTCCGGGAGCGGGGTGTCCAGCTAACAGCGGCTCAAGAACTAATGATC 180
DB 208 SerGlnGlySerGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 227
QY 181 CAGCAGTTGCTGGCGGCCCACTGCAGTCAAAACAGCTCTCTCCGACCCAGCCCAAA 240
DB 228 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 247
QY 241 GTCACGCCCTGGGCCCTGGCGGAGACCCCACTGCAGTCCCGAGATCCCGCCAGCAACGCTTT 300
DB 248 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 267
QY 301 GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCTGGACTTCGCTAAGCAA 360
DB 268 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 287
QY 361 GTCCTGGTTTCTCAGCTGGCGGAGGACCCAGATCCGCTCTCGAAGGCATCCACT 420
DB 288 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 307
QY 421 ATCAGATCATGTCTAGAGACAGCGCGCTACAAACAGCGCTTCCACCCAGCAGAGTGTATCACC 480
DB 308 IleGluIleMetLeuLeuGluThrAlaArgTyrAsnHisGluThrGluCysIleThr 327
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGGACGACTTCCACCGTCCAGGCTGCGAGGTGGAG 539
DB 328 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 347
QY 540 TTCATCAACCCCATCTTCAGATTCTCGGGGCGCATCGGGGCTGGGGCTGGAGCAGCT 599
DB 348 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 367
QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGCGCCCAACGTCAGGAG 659
DB 368 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 387
QY 660 CCGGGCGGCTGAGGCTTGCAGCAGCCCTACGTGAGGCGCTGCTGCTTACACGCGC 719
DB 388 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 407
QY 720 ATCAAGAGCGCGCAGGACGAGTGGCG 746
DB 408 IleLysArgProGlnAspGlnLeuArg 416
```

RESULT 10

```
US-08-776-844-2
; Sequence 2, Application US/08776844
; Patent No. 6277976
; GENERAL INFORMATION:
; APPLICANT: ENMARK, EVA
; APPLICANT: GUSTAFSSON, JAN
```



```

? APPLICANT: ENMARK, EVA
?           GUSTAFSSON, JAN
? TITLE OF INVENTION: OR-1 ON ORPHAN RECEPTOR BELONGING
? TO THE NUCLEAR RECEPTOR FAMILY
?
? NUMBER OF SEQUENCES: 11
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Banner & Witcoff
? STREET: 1001 G Street, NW
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20001
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq for Windows Version 2.0
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/909,326
? FILING DATE: 19-JUL-2001
? CLASSIFICATION: 435
?
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/776,844
? FILING DATE: 1997-06-24
? APPLICATION NUMBER: UK 9413536.2
? FILING DATE: 16-AUG-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Kagan, Sarah A
? REGISTRATION NUMBER: 32141
? REFERENCE/DOCKET NUMBER: 00487.04029
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-508-9100
? TELEFAX: 202-508-9299
? TELEX: <Unknown>
?
? INFORMATION FOR SEQ ID NO: 2:
?     SEQUENCE CHARACTERISTICS:
?         LENGTH: 446 amino acids
?         TYPE: amino acid
?         STRANDEDNESS: single
?         TOPOLOGY: linear
?     MOLECULE TYPE: protein
?     SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-909-326-2

Alignment Scores:
Pred No.:          9,39e-93          Length:      446
Score:            1101.00             Matches:    225
Percent Similarity: 92.77%             Conservative: 6
Best Local Similarity: 90.36%           Mismatches: 12
Query Match:       80.13%              Indels:     7
DB:                2                  Gaps:       1

US-10-712-629B-2 (1-746) x US-09-909-326-2 (1-446)
Qy   1  AAGATTTCGAAACAGCAGCAGCAGGAGTCACTGCAGTCACAGTCGCAGTCACCTGTGGGGCCG 60
Db   160 LysileGlnLysGlnGlnGln-----GlnProProProThr 173
Qy   61  CAGGGCAGCAGCAGCTACGCTCTCGGGCTTGCGGCTTCCCCTGGTGATCTCAGGCAGGC 120
Db   174 GluProAlaSerGlySerSerAlaAargProAlaAlaSerProGlyThrSerGluAlaSer 193
Qy   121 AGCCAGGGCTCCGGGGAGGCGAGGTGTCACAGCTAACAGCGGCTCACGAACCTAATGATC 180
Db   194 SerGlnGlySerGlyGlnGlyGlnGlnLeuThrAlaAleGlnLeuMetIle 213
Qy   181 CAGCAGTTGTTGGGGGCCCAACTGCAGTGCAACAAACGCTCCTCTCCGACCAGCCCCAAA 240
Db   214 GlnGlnLeuValAlaAlaGlnLeuGlnCysaenLysArgSerPheSerAspGlnProLys 233
Qy   241 GTACAGCCCTGGCCCTGGGCGCAGACCCCCAGTCCCGAGATGCCCGCAGCAACGCTTT 300
Db   234 ValThrProTPpLoeUgLYalaAspProGlnSerArgAspAlaAargGlnGlnArgPhe 253

```

RESULT 12
US-09-909-326-2
; Sequence 2, Application US/090909326
; Patent No. 6617120
; GENERAL INFORMATION:

QY 301 GCCACATTCACGAGCTGGCCATCATCTAGTCCAGAGATCTGGACTTCGTGAAGCAA 360
Db 254 AlaHisPheThrGluLeuAlaIleSerValGlnGluLeuValAspPheAlaLysGln 273
QY 361 GTCCCTGGTTCTGACGTGGCCGGGAGGACAGATCGCCCTCTCTGAAGGCATCCACT 420
Db 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnAlaLeuLeuLysAlaSerThr 293
QY 421 ATCCAGATCATGCTGTAGACAGCCAGCGCTACAAACCACGAGACAGAGTGTATCACC 480
Db 294 IleGluLeuMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGACGACTTCCACCGTCAGCGCTGCAGGTGGAG 539
Db 314 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTTCATCAACCCCATCTCGAGTTCTCGCGGGCCATCGCGGGCTGGCCCTGGACGCGCT 599
Db 334 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 353
QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGGCCGACCGCCCAACGTCGACGAG 659
Db 354 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY 660 CCGGGCGCGTGGAGCGTTGACAGCCCTACGTGGAGCGCTGTCTCTACACGCGC 719
Db 374 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuSerTyrThrArg 393
QY 720 ATCAAGAGCGCGAGGACGCTGGCC 746
Db 394 IleLysArgProGlnAspGlnLeuArg 402

RESULT 13

US-08-372-652-3
; Sequence 3, Application US/08372652
; Patent No. 5932699
; GENERAL INFORMATION:
; APPLICANT: Moore, David
; APPLICANT: Seol, Wongi
; APPLICANT: Choi, Hwang-Sik
; TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING
; TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street, Suite 3100
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/372,652
; FILING DATE: 13-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul I.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/246001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-372-652-3

Alignment Scores:

Pred. No.: 3e-92 Length: 446
Score: 1095.50 Matches: 223
Percent Similarity: 93.57% Conservative: 10
Best Local Similarity: 89.56% Mismatches: 15
Query Match: 79.73% Indels: 2
DB: 1 Gaps: 1

US-10-712-629B-2 (1-746) x US-08-372-652-3 (1-446)

QY 1 AAGATTCGGAACACAGCAGCAGGAGTCACAGTCACAGTCGCAGTCACCTGTGGGGCCG 60
Db 155 GlnIleArgLysArgIleGlnLys---GlnGlnGlnGlnGlnProProSer 173
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCTTCCCTGGTGGATCTGAGGCGAGGC 120
Db 174 GluProAlaIleSerSerGlyArgProAlaIleAsnIlePheSerProGlyThrSerGluAlaSer 193
QY 121 AGCCAGGCTCCGGGGAAGCGAGGTGTCCAGCTAACACGCGCTCAAGAACTAATGATC 180
Db 194 SerGlnGlySerGlyGluGlyIleGlnLeuThrAlaIleGlnGluLeuMetIle 213
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCGAGTCACAAACGCTCTTCTCCGACCGCCCAA 240
Db 214 GlnGlnLeuValAlaIleGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTCAGCCCTGGGCCCTGGCGCAGACCCCGAGTCCCGAGATGCCCGCAGCAACGCTTT 300
Db 234 ValThrProIleProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 253
QY 301 GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGAGAGATCTGAGCTTCGTGAAGCAA 360
Db 254 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
QY 361 GTCCCTGGTTCTGACGTGGCCGGGAGGACAGATCGCCCTCTCTGAAGGCATCCACT 420
Db 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY 421 ATCCAGATCATGCTGTAGACAGCAGCGCTACAAACCACGAGACAGAGTGTATCACC 480
Db 294 IleGluLeuMetLeuLeuGlnThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGACGACTTCCACCGTCAGCGCTGCAGGTGGAG 539
Db 314 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTTCATCAACCCCATCTCGAGTTCTCGCGGGCCATCGCGGGCTGGCCCTGGACGCGCT 599
Db 334 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 353
QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGGCCGACCGCCCAACGTCGACGAG 659
Db 354 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY 660 CCGGGCGCGTGGAGCGTTGACAGCCCTACGTGGAGCGCTGTCTCTACACGCGC 719
Db 374 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuSerTyrThrArg 393
QY 720 ATCAAGAGCGCGAGGACGCTGGCC 746
Db 394 IleLysArgProGlnAspGlnLeuArg 402

RESULT 14

PCT-US95-16311-3
; Sequence 3, Application PC/TUS9516311
; GENERAL INFORMATION:
; APPLICANT: Moore, David
; APPLICANT: Seol, Wongi
; APPLICANT: Choi, Hwang-Sik
; TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING

Alignment Scores:			
Pred. No.:	4, 12e-92	Length:	443
Score:	1094.00	Matches:	224
Percent Similarity:	92.37%	Conservative:	6
Best Local Similarity:	89.96%	Mismatches:	13
Query Match:	79.62%	Indels:	7

```
DB: 1 1 Gaps: 1
US-10-712-629B-2 (1-746) x US-08-342-411A-4 (1-443)
QY 1 AAGATTCGGAAACAGCAGCAGCAGGAGTCACAGTCGCGAGTCACCTGTGGGGCCG 60
Db |||:::|||
157 LysIleGlnLysGlnGlnGln-----GlnProProProThr 170
QY 61 CAGGGCAGCAGCAGCTCTGGGCTTGGGCTTCCCTGGTGGATCTGAGGAGGC 120
Db ::|||::|
171 GluProAlaSerGlySerAlaArgProAlaAlaSerProGlyThrSerGluAlaSer 190
QY 121 AGCAGGGCTCCGGGAGCGAGGTGTCAGCTAACAGCGGCTCAAGAACTAATGATC 180
Db |||:::|||
191 SerGlnGlySerGlyGluGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 210
QY 181 CAGCAGTTGTTGGGGCCCAACTGCAGTCGCAACAAACGCTCTTCTCCGACCAAGCCCAA 240
Db |||:::|||
211 GlnGlnLeuValAlaValGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 230
QY 241 GTCAGGCCCTGGGCCCTGGGCGCAGACCCCGAGTCCCGAGATGCCCGCCAGCAACGCTTT 300
Db |||:::|||
231 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 250
QY 301 GCCCACTTCAGGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
Db |||:::|||
251 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 270
QY 361 GTCGCTGGTTTCTGCGAGTGGGGCGGAGGACCAGATCGCCCTCTGAAGGATCCACT 420
Db |||:::|||
271 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 290
QY 421 ATCGAGATCATGCTGTAGACAGCCAGCGCTTACAAACCAGCAGACAGAGTGTATCAC 480
Db |||:::|||
291 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 310
QY 481 TTCTTGAG-GACTTCACCTACAGCAGCAGACTTCCACCGTCGAGGCCCTGCAGGTGGAG 539
Db |||:::|||
311 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 330
QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGCCATCGCGGCTGGGCTGGAGCGCT 599
Db |||:::|||
331 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 350
QY 600 GAGTACGCCCTGTCTATCGCCATCAACATCTTCTCGGCCGACCGGCCCAACGTCAGGAG 659
Db |||:::|||
351 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 370
QY 660 CCGGGCCGCTGGAGGCGTTGCAGCAGCCCTACGTGGAGGCGCTGTCTACACGGC 719
Db |||:::|||
371 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 390
QY 720 ATCAAGAGGCGCAGGACCACTGCGC 746
Db |||:::|||
391 IleLysArgProGlnAspGlnLeuArg 399
```

Search completed: November 25, 2005, 22:33:18
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2005, 22:23:04 ; Search time 150 Seconds
(without alignments)
4156.009 Million cell updates/sec

Title: US-10-712-629B-2

Perfect score: 1374

Sequence: 1 agattcggaacagcagca.....ggccgcaggaccagctgcgc 746

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DSV=xlp
-Q=/cgn2/USFTO.spool_p/US10712629/runat 25112005 143038 11020/app_query.fasta_1.903
-DB=Published Applications AA Main -QFMT=faetan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAESTZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10712629 @CGN 1 1 307 @runat 25112005 143038 11020 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main.*

1: /cgn2_6/ptodata1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata1/pubpaa/US08_PUBCOMB.pcp.*
3: /cgn2_6/ptodata1/pubpaa/US09_PUBCOMB.pcp.*
4: /cgn2_6/ptodata1/pubpaa/US10A_PUBCOMB.pcp.*
5: /cgn2_6/ptodata1/pubpaa/US10B_PUBCOMB.pcp.*
6: /cgn2_6/ptodata1/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1231	89.6	461	4	US-10-013-823-3
2	1231	89.6	461	4	US-10-018-007-4
3	1231	89.6	461	5	US-10-899-458-12
4	1231	89.6	461	5	US-10-500-912-4
5	1215.5	88.5	460	4	US-10-327-813-4
6	1215.5	88.5	460	4	US-10-329-668-4
7	1215.5	88.5	460	4	US-10-717-049-4
8	1215.5	88.5	460	5	US-10-509-197-4
9	1162	84.6	297	3	US-09-925-297-635
10	1101	80.1	446	3	US-09-909-446-2
11	1101	80.1	446	3	US-09-909-325-2

12	1101	80.1	446	3	US-09-909-326-2	Sequence 2, Appli
13	1098.5	79.9	446	4	US-10-013-823-2	Sequence 2, Appli
14	1098.5	79.9	446	5	US-10-899-458-10	Sequence 10, Appli
15	1094	79.6	443	5	US-10-899-458-8	Sequence 8, Appli
16	1005	73.1	263	4	US-10-418-007-8	Sequence 8, Appli
17	777	56.6	264	4	US-10-043-487-254	Sequence 254, App
18	719.5	52.4	445	5	US-10-899-458-4	Sequence 4, Appli
19	718.5	52.3	445	5	US-10-899-458-2	Sequence 2, Appli
20	707.5	51.5	377	4	US-10-236-417-4	Sequence 4, Appli
21	705.5	51.3	553	4	US-10-338-411-25	Sequence 25, Appli
22	705.5	51.3	553	4	US-10-389-640-25	Sequence 25, Appli
23	704.5	51.3	451	4	US-10-236-417-8	Sequence 8, Appli
24	702.5	51.1	447	4	US-10-327-813-2	Sequence 2, Appli
25	702.5	51.1	447	4	US-10-329-668-2	Sequence 2, Appli
26	702.5	51.1	447	4	US-10-418-007-2	Sequence 2, Appli
27	702.5	51.1	447	4	US-10-429-160-12	Sequence 12, Appli
28	702.5	51.1	447	4	US-10-236-417-2	Sequence 2, Appli
29	702.5	51.1	447	4	US-10-236-417-6	Sequence 6, Appli
30	702.5	51.1	447	4	US-10-717-049-2	Sequence 2, Appli
31	702.5	51.1	447	5	US-10-899-458-6	Sequence 6, Appli
32	702.5	51.1	447	5	US-10-921-023-2	Sequence 2, Appli
33	702.5	51.1	447	5	US-10-500-912-2	Sequence 2, Appli
34	702.5	51.1	447	5	US-10-509-197-2	Sequence 2, Appli
35	692	50.4	433	4	US-10-278-945-2	Sequence 2, Appli
36	691	50.3	296	5	US-10-921-023-39	Sequence 39, Appli
37	680	49.5	191	4	US-10-104-047-2537	Sequence 2537, Ap
38	679	49.4	220	4	US-10-418-007-6	Sequence 6, Appli
39	660.5	48.1	511	5	US-10-921-023-4	Sequence 4, Appli
40	607.5	44.2	409	5	US-10-921-023-6	Sequence 6, Appli
41	504	36.7	136	5	US-10-921-023-27	Sequence 27, Appli
42	462	33.6	200	5	US-10-921-023-28	Sequence 28, Appli
43	400.5	29.1	316	4	US-10-468-199-74	Sequence 74, Appli
44	383	27.9	66	3	US-09-864-761-48359	Sequence 48359, A
45	367	26.7	167	4	US-10-424-599-177842	Sequence 177842,

ALIGNMENTS

RESULT 1
US-10-013-823-3
; Sequence 3, Application US/10013823
; Publication No. US20020116731A1
; GENERAL INFORMATION:
; APPLICANT: Guenther, Catherine
; APPLICANT: Phillips, Russell
; APPLICANT: Allen, Keith D.
; APPLICANT: Zhang, Qin
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING RETINOID X
; TITLE OF INVENTION: RECEPTOR INTERACTING PROTEIN GENE DISRUPTIONS
; FILE REFERENCE: R-684
; CURRENT APPLICATION NUMBER: US/10/013,823
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/254,801
; PRIOR FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 60/309,404
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-013-823-3

Alignment Scores:
Score: 7.69e-88 Length: 461
Seq. No.: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 4 Gaps: 0

[illegible]

[illegible][illegible]

Qy	481	TTCTTGAG - GACTTCACCTACAGCAGACGACCTTCACCGTGCAGGCGCTGCAGGTGGAG	539
Db	329	PheLeuLyAspPheThrTySerLyAspAspPheHisArgAlaGlyLeuGlnValGlu	348
Qy	540	TTTCATCAACCCCATCTTCGAGTTCTCGCGGGCCATCGCGGGCTGGCGCTGCAGACGCT	599
Db	349	PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla	368
Qy	600	GAGTACGCGCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGGCCCAACGTGCAGGAG	659
Db	369	GluTyAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu	388
Qy	660	CGCGGCGCGGTGGAGCGTTGCAGCAGCGCTTACGTGGAGGCGCTGCTGCTCCTACACGCG	719
Db	389	ProGlyArgValGluAlaLeuGlnProTyValGluAlaLeuSerTyThrArg	408
Qy	720	ATCAGAGGCGCAGGACCGCTGCGC	746
Db	409	IleLyAsArgProGlnAspGlnLeuArg	417

RESULT 5

```

US-10-327-813-4
; Sequence 4, Application US/10327813
; Publication No. US20030181420A1
; GENERAL INFORMATION:
; APPLICANT: Bayne Christopher D.
; APPLICANT: Johnson Alan T.
; APPLICANT: Lu Shao-Po
; APPLICANT: Mohan Raju
; APPLICANT: Griffith Ronald C.
; TITLE OF INVENTION: Modulators Of LXR
; FILE REFERENCE: 38205-3005
; CURRENT APPLICATION NUMBER: US/10/327,813
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/342,707
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-327-813-4

```

Alignment Scores:		
Pred. NO.:	1.26e-86	460
Score:	1215.50	247
Percent Similarity:	99.20%	Conservative: 0
Best Local Similarity:	99.20%	Mismatches: 1
Query Match:	88.46%	Indels: 2
DB:	4	Gaps: 1

US-10-712-629B-2 (1-746) x US-10-327-813-4 (1-460)

[illegible]

301	Qy	GCCTCATCTTCAGGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCCTGCTAAGCAA	366
268	Db	AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaIlysGln	287
361	Qy	GTGGCTCGTTTCTCGAGACTGGGGCGGGAGGACAGATCGCCCTCTGAAGGCATCCACT	420
288	Db	ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuIysAlaSerThr	307
421	Qy	ATCGAGATCATGTGCTAGAGACAGCAGCGCGCTTACAACCAACGAGACAGAGTGTATCACC	480
308	Db	IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr	327
481	Qy	TTCTTTGAG-GACTTTCACCTACAGCAAGGACGACTTCCACCGTGCAGGGCTCGCAGGTGGAG	539
328	Db	PheLeuIysAspPheThrTyrSerTyrLeuAspAspPheHisArgAlaGlyLeuGlnValGlu	347
540	Qy	TTTCATCAACCCCATCTTCGAGTTCTTCGGGGCCCATGCGCGCGCTGGGGCTCGACACGCT	599
348	Db	PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla	367
600	Qy	GAGTACGCCCTGTCTATCGCCATCAACNTCTTCTCGGCGGACCGGCCCAACGTGCAGGAG	659
368	Db	GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu	387
660	Qy	CCGGGCGCGTGGAGGGTGTTCAGCAGCGCCCTACGTGGAGGGGCTCGTCTCTACACGGCG	719
388	Db	ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg	407
720	Qy	ATCAAGAGGCGCAGGACCACTGGCG	746
408	Db	IleIysArgProGlnAspGlnLeuArg	416

RESULT 6

```

US-10-329-668-4
; Sequence 4, Application US/10329668
; Publication No. US2003021211A1
; GENERAL INFORMATION:
; APPLICANT: Martin Richard
; APPLICANT: Brenton Todd
; APPLICANT: Kahl Jeffrey Dean
; APPLICANT: Wang Tie-Lin
; TITLE OF INVENTION: HETEROCYCLIC MODULATORS OF NUCLEAR RECEPTORS

```

```

; FILE REFERENCES: 38205-3001
; CURRENT APPLICATION NUMBER: US/10/329,668
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/342,720
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-329-668-4

Alignment Scores:
Pred. No.: 1,26e-86
Score: 1215.50
Percent Similarity: 99.20%
Best Local Similarity: 99.20%
Query Match: 88.46%
DP: 4

Length: 460
Matches: 247
Conservative: 0
Mismatch: 1
Indels: 1
Gaps: 1

```

US-10-712-629B-2 (1-746) x US-10-329-668-4 (1-460)

QY	1	AAGATTTCGGAAACAGCAGCAGAGAGTACAGTCCAGTCCAGTCACTCTGGGCGC	60
Db	169	LysIleAtrGlys---GlnGlnGlnGlnSerGlnSerProValGlyPro	187
QY	61	CAGGCGCAGCAGCTCAGCCCTCTGGGCGCTGCCCTGGTGATCTCAGGCGAGCG	120
Db	188	GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly	207

Qy	121	AGCCAGGCTCCGGGAGGCGAGGGTGTCAGCTTAACAGCGCTCAAGAACTAATGATC	180
Db	208	SerGInGlySerGlyGluGlyGluGlyValGInLeuThrAlaAlaGInGluLeuMetIle	227
Qy	181	CAGCAGTTGGTGGCGGCCCAACTGCAGTGTCAACAAACGCTCTCTTCCACACAGCCCAAA	240
Db	228	GInGInLeuValAlaAlaGInLeuGInCysAenLysArgSerPheSerAspGInProLys	247
Qy	241	GTCAGCCCTGGCCCTGGCGGCAGACCCCAAGTCCCGAGATGCCCGCCAGACACGCTTT	300
Db	248	ValThrProThrProLeuGlyAlaAspProGInSerArgAspAlaArgGInGInArgPhe	267
Qy	301	GCCCACTTCACGGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACATTCGTAAACAA	360
Db	268	AlaHisPheThrGluLeuAlaIleIleSerValGInGluIleValAspPheAlaLysGIn	287
Qy	361	GTGCTGTGTTCTCCAGCTGGCGCGGAGGACCAAGATCGCCCTCTCGAAGGATCCACT	420
Db	288	ValProGlyPheLeuGInLeuGlyArgGluAspGInIleAlaLeuLeuLysAlaSerThr	307
Qy	421	ATCAGAGTCATGCTCTTAGAGACACGACGAGCGCTACAACACGAGACAGAGTGATCACCC	480
Db	308	IleGluIleMetLeuLeuGluThrIleArgAlaGlyTyrAenHisGluThrGluCysIleThr	327
Qy	481	TTCTTGAG - GACTTCACCTACAGCAAGGACGACTTCACCGTCGAGGCGTGCAGGTGGAG	539
Db	328	PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGInValGlu	347
Qy	540	TTCATCAACCCCATCTTCGAGTTCTCGGGGCCATGCGCGCGCTGGCGCTGCACGACGCT	599
Db	348	PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla	367
Qy	600	GAGTACGCCCTGCTCATCGCCATCAACTCTTCTCGGCGGACCGGCCCAACGTGCAGGAG	659
Db	368	GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGInGlu	387
Qy	660	CCGGGCGCGTGGAGCGGTTCGACGAGCCCTACGTGGAGGCGCTGTCCTACACGCGC	719
Db	388	ProGlyArgValGluAlaLeuGInProTyrValGluAlaLeuLeuSerTyrThrArg	407
Qy	720	ATCAAGAGCGCGAGGACCAAGCTGCGC	746
Db	408	IleLysArgProGInAspGInLeuArg	416

RESULT 7

US-10-717-049-4

00 10 717 043-1
; Sequence 4, Application US/10717049

Publication No. US20040180942A1

GENERAL INFORMATION:

APPLICANT: Martin Richard

APPLICANT: Flatt Brenton Todd

APPLICANT: Kahl Jeffrey Dean

; APPLICANT: Wang Tie-Lin

; TITLE OF INVENTION: HETEROCYCLIC MODULATORS OF NUCLEAR RECEPTORS

; FILE REFERENCE: 38205-3001B

; CURRENT APPLICATION NUMBER: US/10/717,049

; CURRENT FILING DATE: 2003-11-18

; PRIOR APPLICATION NUMBER: 10/329,668

;; PRIOR FILING DATE: 2002-12-20

;; PRIOR APPLICATION NUMBER: 60/342,720

PRIOR FILING DATE: 2001-12-21

; NUMBER OF SEQ ID NOS: 18
: SOFTWARE: Easycio for Win

```

; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0
; SEQ ID NO 4

```

```
; SEQ ID NO 4
: LENGTH: 460
```

```

:      LENGTH: 460
:      TYPE: DPT

```

LIFE: FRI
 ORGANISM:

US-10-717-049-4

4-510-17-07 00

Alignment Scores:

Pred. No.: 1.26e-86

Score: 1215.50

Percent Similarity:	99.20%	Conservative:	0
Best Local Similarity:	99.20%	Mismatches:	1
Query Match:	88.46%	Indels:	2
DB:	4	Gaps:	1

US-10-712-629B-2 (1-746) x US-10-717-049-4 (1-460)

Qy		1	AAGATTCGGAAACAGCAGCAGGAGTCA	CAGTCCAGTCCAGTCTGTGGGCCG	60
Db		169	LysileArglys---GlnGlnGlnSerGln	SerGlnSerProValGlyPro	187
Qy		61	CAGGGCAGCAGCTCAGCCCTCGGGCTT	CCCCCTGGTGATCTCAGGCAGGC	120
Db		188	GlnGlySerSerSerAlaSerGlyProGly	AlaSerProGlyGlySerGluAlaGly	207
Qy		121	AGCCAGGGCTCCGGGGGAAGCGAGGTG	TCCAGCTTAACAGCGGCTCAAGAACA	180
Db		208	SerGlnGlySerGlyGluGlyGluGlyVal	GlnLeuThrAlaAaGlnGluLeuMetIle	227
Qy		181	CAGCAGTTGGTGGCGGCCAACCTCGAGT	GCAGCAACAACTCCTTCTCGACAGCCC	240
Db		228	GlnGlnLeuValAlaAlaGlnLeuGlnCy	sAsnLyArGserPheSerAspGlnProLy	247
Qy		241	GTCAAGCCCTGGCCCTCGGGCGCAGACC	CCCCAGTCCCGAGATGCCCGCACAA	300
Db		248	ValThrProTrpProLeuGlyAlaAsppro	GlnSerArgAspAlaArgGlnArgPhe	267
Qy		301	GCCCACTTCACGGAGCTGGCCATCATCT	CAGTCCAGGAGATCGTGGACTTCGCT	360
Db		268	AlaHisphethrGluLeuAlaIleIleSer	ValGlnGluIleValAspPhealalyGln	287
Qy		361	GTGCCTGGTTCTCGAGCTGGCGCGGAG	GCACAGATCGGCCTCCTGAAGGCATCCA	420
Db		288	ValProglyPheLeuGlnLeuGlyArgGlu	AspGlnIleAlaLeuLeulyAlaSerThr	307
Qy		421	ATCGAGATCATGTGTAGACACAGCCAG	CGCGCTACACACAGCAGCAGAGTGATCAC	480
Db		308	IleGluIleMetLeuLeuGluThrAlaarg	ArgTyrAenHisGluThrGluCysIlethr	327
Qy		481	TTCTTGAG-GACTTCACCTACAGCAAG	GCACGACTTCCACCGTGAGGCCTCGAC	539
Db		328	PheLeulyAspPheThrTyrSerLysAsp	AspPheHisArgAlaGlyLeuGlnValGlu	347
Qy		540	TTCAATCAACCCCATCTTCGAGTTCTCG	GGGCCCATGCGCGCGCTGGCGCTGAC	599
Db		348	PheIleAsnProIlePheGluPheSerArg	AlaMetArgLeuGlyLeuAspAspAla	367
Qy		600	GAGTACGCCCTGCTCATCGCCATCAAC	ATCTTCTCGGCCGACCGGCCCAACGTG	659
Db		368	GluTyrAlaLeuLeuIleAlaIleAsnIle	PheSerAlaAspArgProAsnValGlnGlu	387
Qy		660	CCGGCGCGCTGGAGGGCTTGACAGCAG	CCCTTACCTGGAGGGCTGCTCTCTAC	719
Db		388	ProGlyArgValGluAlaLeuGlnGlnPro	TyrValGluAlaLeuLeuSerTyrThrArg	407
Qy		720	ATCAAGAGCGCAGGACCACTGCGC		746
Db		408	IleLysArgProGlnAspGlnLeuArg		416

RESULT 8

US-10-509-197-4

; Sequence 4, Application US/10509197

Publication No. US20050171084A1

GENERAL INFORMATION:

; APPLICANT: Burbidge, Stephen A.

; APPLICANT: Cairns, William J.

; APPLICANT: Irving, Elaine A

; APPLICANT: Parsons, Andrew

; APPLICANT: Richardson, Jill C.

APPLICANT: Soden, Peter E.

; APPLICANT: Vinson, Mary

APPLICANT: Watson, Mike A.

APPLICANT: WHITNEY, KARL D.

```
; TITLE OF INVENTION: Methods of Treatment with LXR Modulators
; FILE REFERENCE: P51332
; CURRENT APPLICATION NUMBER: US/10/509,197
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: PCT/US03/09225
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: 60/368,424
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-509-197-4

Alignment Scores:
Pred. No.: 1,26e-86 Length: 460
Score: 1215.50 Matches: 247
Percent Similarity: 99.20% Conservative: 0
Best Local Similarity: 99.20% Mismatches: 1
Query Match: 88.46% Indels: 2
DB: 5 Gaps: 1

US-10-712-629B-2 (1-746) x US-10-509-197-4 (1-460)

Qy 1 AAGATTGCGAAACAGCAGCAGGAGTCACAGTCAGTCGCGAGTCACCTGTGGGGCCG 60
Db 169 Lyslearglys--GlnGlnGlnSerGlnSerGlnSerGlnSerProValGlyPro 187

Qy 61 CAGGGCAGCAGCAGCTCAGCCTCGGGCTGGGGCTTCCCTGGTGGATCTGAGGAGGC 120
Db 188 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 207

Qy 121 AGCCAGGCTCCGGGAGCGAGGTGTCAGCTACACGGCTCAAGAACTAATGATC 180
Db 208 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 227

Qy 181 CAGCAGTGTGTGGCGGCCCAACTGCAGTGCAACAAAGCTCTTCTCCGACCAGCCCAA 240
Db 228 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 247

Qy 241 GTCAGCGCTTGGCGGGCGAGACCCCGAGTCCCGAGATGCCGCCAGCAGCGTTT 300
Db 248 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 267

Qy 301 GCCACATTCAGGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
Db 268 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluLeuValAspPheAlaLysGln 287

Qy 361 GTGCTGTGTTTCTGACAGCTGGGCGGAGGACCAGATCGCCCTCTCTGAAGGATCCACT 420
Db 288 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 307

Qy 421 ATCAGATCATGTGCTAGACAGCAGCGGCTACAAACCAGCAGACAGAGTGTATCACC 480
Db 308 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 327

Qy 481 TTCTTGAG-GACTTTCACCTACAGCAAGCAGACTTCCACCGTCAGGCTCGAGTGGAG 539
Db 328 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 347

Qy 540 TTCATCAACCCCATCTTTCGAGTTCTCGGGGCCATCGCGGCGTGGCCCTGGACGACGCT 599
Db 348 PheIleAsnProIlePhePhePheSerArgAlaMetArgLeuGlyLeuAspAspAla 367

Qy 600 GAGTACGCCCTGTCTATCGCCATCAACATCTTCTCGGCCGACCGGCCCAACGTCGAGAG 659
Db 368 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 387

Qy 660 CCGGGCCGCTGGAGGCGTTGACAGCCCTACGCTGGAGGCGCTGCTCTACACCGCG 719
Db 388 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 407
```

```
Qy 720 ATCAAGAGCCCGCAGGACGACGAGTCGCGC 746
Db 408 IleLysArgProGlnAspGlnLeuArg 416

RESULT 9
US-09-925-297-635
; Sequence 635, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 635
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (222)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (242)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (254)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (268)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (269)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (274)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (280)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (282)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (295)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (296)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-635

Alignment Scores:
Pred. No.: 1,83e-82 Length: 297
Score: 1162.00 Matches: 242
Percent Similarity: 97.19% Conservative: 0
Best Local Similarity: 97.19% Mismatches: 6
Query Match: 84.57% Indels: 3
DB: 3 Gaps: 0

US-10-712-629B-2 (1-746) x US-09-925-297-635 (1-297)

Qy 3 GATTTCGGAACACGACGACGAGGAGTCAGTCAGTCAGTCACCTGTGGGGCGCA 62
Db 8 AspSerGluThrThrAlaAlaGlyValThrValThrValAlaValThrCysGlyAlaAla 27

Qy 63 GGCACGACGAGCTCAGCCTCTGGGCTTCCCTCTGGTGGATCTCTGAGCGAGCAG 122
```



```
Db 28 -GlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGlySe 47
QY 123 CCAGGGCTCCGGGAGCGAGGGTGTCCAGCTAAACAGCGGCTCAAGAACTAATGATCCA 182
Db 47 rGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIleGl 67
QY 183 GCAGTTGGTGGCGGCCCACTGCAGTGCACAAACGCTCTCTCCGACCCAGCCCAAGT 242
Db 67 nGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLysVa 87
QY 243 CAGCCCTGGCCCTGGGCGCAGACCCAGTCCGAGATCCCGCCGCGCAGCAACGCTTTGC 302
Db 87 lThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPheAl 107
QY 303 CCACTTCACGGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAAGT 362
Db 107 aHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGlnVa 127
QY 363 GCTGTGTTTCTCGAGCTGGCGCGGAGGACCAAGATCGCCCTCTGAAAGCATCCACTAT 422
Db 127 lProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThrIl 147
QY 423 CGAGATCATGCTCTAGAGACACCCAGCGGCTCAACACCGAGACAGAGTGATACCTT 482
Db 147 eGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThrPh 167
QY 483 CTTGAG-GACTTCACCTACAGCAAGGACGACTTCCACCGTGACGGCTCGAGTGCGATT 541
Db 167 eLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGluPh 187
QY 542 CATCAACCCCATCTTCGAGTTCTCGCGGGCCATCGCGCGCTGGGCTGGACGACGCTGA 601
Db 187 eIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAlaGl 207
QY 602 GTACGCCCTGCTCATCGCCATCAACATCTTCTCGCGCGACCCGCCCAACGTGCGAGGCC 661
Db 207 uTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArg***AsnValGlnGluPr 227
QY 662 GGGCGCGGTGGAGCGCTTCAGCAGCGCCTACGT-GGAGGCGCTGCTCTACACCGCA 720
Db 227 oGlyArgValGluAlaLeuGlnGlnProTyrValGlyGlyAla***ValLeuHisAlaHl 247
QY : 721 TCAAGAGCGCGCAGGACCAAGCTCGC 745
Db 247 sGlnGluAlaAlaGlyPro***Ala 255
```

RESULT 10

US-09-909-446-2

; Sequence 2, Application US/0909446

; Patent No. US20020052489A1

; GENERAL INFORMATION:

; APPLICANT: ENMARK, EVA

; TITLE OF INVENTION: OR-1 ON ORPHAN RECEPTOR BELONGING

; TO THE NUCLEAR RECEPTOR FAMILY

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner & Witcoff

; STREET: 1001 G Street, NW

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/909,446

; FILING DATE: 19-Jul-2001

; CLASSIFICATION: <Unknown>

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/776,844
; FILING DATE: <Unknown>
; APPLICATION NUMBER: UK 9413536.2
; FILING DATE: 16-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32141
; REFERENCE/DOCKET NUMBER: 00487.04029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-909-446-2
```

Alignment Scores:

Pred. No.:	1.18e-77	Length:	446
Score:	1101.00	Matches:	225
Percent Similarity:	92.77%	Conservative:	6
Best Local Similarity:	90.36%	Mismatches:	12
Query Match:	80.13%	Indels:	7
DB:	3	Gaps:	1

US-10-712-629B-2 (1-746) x US-09-909-446-2 (1-446)

```
QY 1 AGATTCCGAAACAGCAGCAGCAGGAGTCACAGTCACAGTCGCGAGTCACCTGTGGGCGG 60
Db 160 LysIleGlnLysGlnGlnGln-----GlnProProProThr 173
QY 61 CAGGGCAGCAGCAGCTCAGCCCTCGGGCTTGGGGCTTCCCTGGTGATCTGAGGCGAGC 120
Db 174 GluProAlaSerGlySerSerAlaArgProAlaAlaSerProGlyThrSerGluAlaSer 193
QY 121 AGCCAGGGTCCGGGAGCGAGGGTGTCCAGCTAACAGCGGCTCAAGAACTAATGATC 180
Db 194 SerGlnGlySerGlyGlyGlyGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 213
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACGCTCTTCTCCGACCGCCCAAA 240
Db 214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTCAGCCCTGGCCCTGGGCGCAGACCCCGAGATGCCCGCAGCAACGCTTT 300
Db 234 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 253
QY 301 GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 254 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
QY 361 GTGCGCTGGTTTCTCGAGCTGGCGGGAGGAGCAGATCGCCCTCTCTGAAGGATCACT 420
Db 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY 421 ATCGAGATCATGCTGTAGAGACAGCGCGCTACAACACGAGACAGAGTGTATCACC 480
Db 294 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGGACGACTTCCACCGTCGAGGCTCGCAGGTGGAG 539
Db 314 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGGCGCATCGCGGCTGGGCTGGACGCGCT 599
Db 334 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAla 353
```


APPLICATION NUMBER: US/09/909,326
FILING DATE: 19-Jul-2001
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/776,844
FILING DATE: 1997-06-24
APPLICATION NUMBER: UK 9413536.2
FILING DATE: 16-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32141
REFERENCE/DOCKET NUMBER: 00487.04029
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 446 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-909-326-2

Alignment Scores:
Pred. No.: 1,18e-77 Length: 446
Score: 1101.00 Matches: 225
Percent Similarity: 92.77% Conservative: 6
Best Local Similarity: 90.36% Mismatches: 12
Query Match: 80.13% Indels: 7
DB: 3 Gaps: 1

US-10-712-629B-2 (1-746) x US-09-909-326-2 (1-446)

```
QY 1 AAGATTTCGAAACAGCAGCAGCAGGAGTCACAGTCACAGTCGCGAGTCACCTGTGGGGCGG 60
Db 160 LyslleGlnLysGlnGlnGln-----GlnProProProThr 173
QY 61 CAGGCGCAGCAGCAGCTCAGCCTCTGGGCTTGGGCTTCCCTGGTGATCTGAGGAGCG 120
Db 174 GluProAlaSerGlySerAlaArgProAlaAlaSerProGlyThrSerGluAlaSer 193
QY 121 AGCCAGGGCTCCGGGGAAGCAGGGGTGTCAGTAAACAGCGCTCAAGAACTAATGATC 180
Db 194 SerGlnGlySerGlyGlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 213
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACGCTCTCTCCGACCGCCCAAA 240
Db 214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTACGCGCTTGGCCCTGGCGGAGACCCCGAGTCCGAGATGCGCGCAGCAGACGCTTT 300
Db 234 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 253
QY 301 GCCCACTTCACGAGCTGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 254 AlaHisPheThrGlnLeuAlaIleIleSerValGlnGlnIleValAspPheAlaLysGln 273
QY 361 GTGCTGTTCTTCAGCTGGCGGAGGAGCAGATCGCCCTCTCTGAAGGCATCCACT 420
Db 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY 421 ATCGAGATCATGCTAGACAGCAGCGCGTACAAACAGCAGACAGAGTGATATCAC 480
Db 294 IleGluIleMetLeuLeuGlnThrAlaArgArgTyrAsnHisGlnThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTCACTACAGCAGGAGCACTTCCACCGTCCAGGCGCTGAGGTGGAG 539
Db 314 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTCATCAACCCCATCTTCAGATTCTCGCGGGCCATCGCGGGCTGGGCGCTGACGACGCT 599
```

```
Db 334 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 353
QY 600 GAGTAGCCCTGCTCATCGCCATCAACATCTTCTCGCGCCGACCGCCCAACGTCGAGAG 659
Db 354 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY 660 CCGGGCGGCTGGAGGCGTTGAGCAGCGCTTACGTCGAGCGCTGCTGTCTTACACGCGC 719
Db 374 ProSerArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 393
QY 720 ATCAAGAGCCGCGCAGCAGCAGCTGCGC 746
Db 394 IleLysArgProGlnAspGlnLeuArg 402
RESULT 13
US-10-013-823-2
; Sequence 2, Application US/10013823
; Publication No. US20020116731A1
; GENERAL INFORMATION:
; APPLICANT: Guenther, Catherine
; APPLICANT: Phillips, Russell
; APPLICANT: Allen, Keith D.
; APPLICANT: Zhang, Qin
; APPLICANT: Baribault, Helene
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING RETINOID X
; TITLE OF INVENTION: RECEPTOR INTERACTING PROTEIN GENE DISRUPTIONS
; FILE REFERENCE: R-684
; CURRENT APPLICATION NUMBER: US/10/013,823
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/254,801
; PRIOR FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 60/309,404
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-013-823-2
Alignment Scores:
Pred. No.: 1,85e-77 Length: 446
Score: 1098.50 Matches: 224
Percent Similarity: 93.57% Conservative: 9
Best Local Similarity: 89.96% Mismatches: 15
Query Match: 79.95% Indels: 2
DB: 4 Gaps: 1
US-10-712-629B-2 (1-746) x US-10-013-823-2 (1-446)
QY 1 AAGATTTCGAAACAGCAGCAGCAGGAGTCACAGTCACAGTCGCGAGTCACCTGTGGGGCGG 60
Db 155 GlnIleArgLysLysArgIleGlnLys---GlnGlnGlnGlnGlnProProProSer 173
QY 61 CAGGCGCAGCAGCAGCTCAGCCTCTGGGCTTGGGCTTCCCTGGTGATCTGAGGAGCG 120
Db 174 GluProAlaAlaSerSerGlyArgProAlaAlaSerProGlyThrSerGluAlaSer 193
QY 121 AGCCAGGGCTCCGGGGAAGCAGGGGTGTCAGTAAACAGCGCTCAAGAACTAATGATC 180
Db 194 SerGlnGlySerGlyGlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 213
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACGCTCTCTTCTCCGACCGCCCAAA 240
Db 214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTACGCGCTTGGCCCTGGCGGAGACCCCGAGTCCGAGATGCGCGCAGCAGACGCTTT 300
Db 234 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 253
QY 301 GCCCACTTCACGAGCTGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
```



```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-899-458-8
```

```
Alignment Scores:
Pred. No.: 4.17e-77 Length: 443
Score: 1094.00 Matches: 224
Percent Similarity: 92.37% Conservative: 6
Best Local Similarity: 89.96% Mismatches: 13
Query Match: 79.62% Indels: 7
DB: 5 Gaps: 1
```

US-10-712-629B-2 (1-746) x US-10-899-458-8 (1-443)

```
QY 1 AGATTTCGGAACAGCAGCAGCAGGAGTCACTCAGTCGCGAGTCACCTGTGGGGCCG 60
Db |||||:|||||||
157 LysIleGlnLysGlnGlnGln-----GlnProProProThr 170
QY 61 CAGGGCAGCAGCAGCAGCTCGCCTCGGCCCTGGGGCTTCCCTGGTGATCTGAGGCAGC 120
Db |||||:|||||||
171 GluProAlaSerGlySerAlaArgProAlaAlaSerProGlyThrSerGluAlaSer 190
QY 121 AGCCAGGGCTCCGGGAAGCGAGGTGTCCAGCTAACACGCGCTCAAGAACTAATGATC 180
Db |||||:|||||||
191 SerGlnGlySerGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 210
QY 181 CAGCACTTGGTGGCGGCCCAACTGCAGTGCACAAACAGTCTCTTCCGACCAGCCCAA 240
Db |||||:|||||||
211 GlnGlnLeuValAlaValGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 230
QY 241 GTCACGCCCTGGCCCTGGCGCAGACCCCGAGTCCGAGATGCCCGCAGCAACGCTTT 300
Db |||||:|||||||
231 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 250
QY 301 GCCCACTTCACGAGGTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db |||||:|||||||
251 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 270
QY 361 GTGCTGGTTCTGTCAGTGGCGCGGAGGACCAAGATCGCCCTCTGAAGGCATCCACT 420
Db |||||:|||||||
271 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 290
QY 421 ATCAGATCATGCTGTAGACAGCAGCGCGCTACACCACGACAGACAGAGTGTATCACC 480
Db |||||:|||||||
291 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 310
QY 481 TTCCTTGAG-GACTTCACTACAGCAAGGACGACTTCCACGTCGACGCTGCAAGGTGGAG 539
Db |||||:|||||||
311 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 330
QY 540 TTCATCAACCCCATCTTCAGATTCTCGCGGCCCATCGCGCGCTGGCGCTGGACAGCGCT 599
Db |||||:|||||||
331 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla 350
QY 600 GAGTAGCCCTGCTCATCGCCATCAACATCTTCTCGGCCGACCGGCCCAACGTGCGAGGAG 659
Db |||||:|||||||
351 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 370
QY 660 CCGGGCGCGGTGGAGCGTTGCAGCAGCCCTACGTGGAGCGCTGTCTCTACACGGCCG 719
Db |||||:|||||||
371 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 390
QY 720 ATCAAGAGCCGCGAGCCAGCTGCGC 746
Db |||||:|||||||
391 IleLysArgProGlnAspGlnLeuArg 399
```

Search completed: November 25, 2005, 22:38:27
Job time : 161 secs

This Page Blank (end)

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (Copy)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2005, 22:17:17 ; Search time 4.5 Seconds
(without alignments)
1004.935 Million cell updates/sec

Title: US-10-712-629B-2

Perfect score: 1374

Sequence: 1 aagattcggaacacgagca.....ggccgagaccagctggcg 746

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 17545 seqs, 3030971 residues

Total number of hits satisfying chosen parameters: 35090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlp

-MODEL=frame_n2p.model -DEV=xlp
-Q=/cn2_1/USFTO_spool_p/US10712629/runat_25112005_143038_11026/app_query.fasta_1.903
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10712629 @CN 1.1 @runat_25112005_143038_11026
-NCPUL6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLG=6 -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Databases:

Published Applications AA New.*
1: /cn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
2: /cn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
3: /cn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
4: /cn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
5: /cn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
6: /cn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
7: /cn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pap.*
8: /cn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	174.5	12.7	505	1	US-10-519-447-4
2	140.5	10.2	1467	1	US-10-821-234-1096
3	117.5	8.6	614	7	US-11-015-546A-20
4	117	8.5	1970	1	US-10-821-234-1641
5	116.5	8.5	1366	1	US-10-821-234-1431
6	113	8.2	735	7	US-11-053-100-44
7	112	8.2	1166	1	US-10-821-234-964
8	111	8.1	624	7	US-11-053-100-48
9	111	8.1	774	7	US-11-053-100-49
10	111	8.1	1560	7	US-11-059-982-1

11	108	7.9	1874	1	US-10-821-234-1182
12	107.5	7.8	1225	7	US-11-053-100-50
13	105.5	7.7	775	7	US-11-053-100-51
14	103.5	7.5	308	1	US-10-967-527A-30
15	103.5	7.5	1366	1	US-10-821-234-1431
16	103	7.5	915	1	US-10-821-234-1514
17	102	7.4	138	1	US-10-667-295-176
18	100.5	7.3	614	7	US-11-015-546A-20
19	99.5	7.2	594	1	US-10-131-826A-10
20	99.5	7.2	558	1	US-10-821-234-921
21	99.5	7.2	1467	1	US-10-821-234-1096
22	98	7.1	757	7	US-11-053-100-47
23	98	7.1	1076	1	US-10-131-826A-219
24	97.5	7.1	888	1	US-10-131-826A-544
25	96.5	7.0	766	1	US-10-821-234-1691
26	96	6.9	239	1	US-10-821-234-1186
27	96	7.0	1311	1	US-10-509-424-5
28	96	6.9	1493	7	US-11-004-057-4
29	95.5	7.0	736	7	US-11-053-100-45
30	95.5	7.0	1186	7	US-11-053-100-46
31	95	6.9	280	1	US-10-821-234-1300
32	95	6.9	1493	7	US-11-004-057-4
33	95	6.9	1532	1	US-10-821-234-914
34	94	6.8	575	1	US-10-131-826A-128
35	93	6.8	405	1	US-10-821-234-1357
36	93	6.8	419	1	US-10-821-234-1556
37	93	6.7	594	1	US-10-131-826A-10
38	93	6.7	1493	7	US-11-004-057-21
39	92.5	6.7	479	1	US-10-821-234-871
40	92.5	6.7	543	1	US-10-689-742-78
41	91.5	6.7	1133	1	US-10-821-234-1219
42	91	6.6	1377	1	US-10-821-234-1070
43	91	6.6	1874	1	US-10-821-234-1182
44	90.5	6.6	503	7	US-11-013-247A-4
45	90.5	6.6	1493	7	US-11-004-057-21

ALIGNMENTS

RESULT 1

US-10-519-447-4

; Sequence 4, Application US/10519447

; Publication No. US20050244829A1

; GENERAL INFORMATION:

; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.

; APPLICANT: Makoto OGINO

; APPLICANT: Hideki ENDOH

; TITLE OF INVENTION: METHOD FOR SCREENING AN AGENT FOR IMPROVING INSULIN RESISTANCE

; FILE REFERENCE: Q85576

; CURRENT APPLICATION NUMBER: US/10/519,447

; CURRENT FILING DATE: 2004-12-30

; PRIOR APPLICATION NUMBER: PCT/JP03/08367

; PRIOR FILING DATE: 2003-07-01

; PRIOR APPLICATION NUMBER: JP 2002-193814

; PRIOR FILING DATE: 2002-07-02

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 505

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-519-447-4

Alignment Scores:

Pred. No.: 9.17e-07 Length: 505

Score: 174.50 Matches: 52

Percent Similarity: 47.13% Conservative: 30

Best Local Similarity: 29.89% Mismatches: 57

Query Match: 12.70% Indels: 35

DB: 1 Gaps: 4

US-10-712-629B-2 (1-746) x US-10-519-447-4 (1-505)

```
QY 295 CCCTTTGCCCACTTCACGAGCTG----- 318
Db 291 LysPheLysHisIleThrProLeuGlnGlnSerLysGluValAlaIleArgIlePhe 310
QY 319 -----CCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 311 GlnGlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGluTyrAlaLysSer 330
QY 361 GTGCGCTGGTTCTCGAGCTGGCGGAGGAGCACAGATCGCCCTCTGAGAGGCATCCACT 420
Db 331 IleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyrGlyVal 350
QY 421 ATCGAGATCATGCTGTAGACAGACGCGCGCTACAACACGAGACAGAGTGATATCACC 480
Db 351 HisGluIleIle-----Tyr-ThrMetLeuAlaSerLeuMetAs 363
QY 481 TTCTTGAGCACTTCACGACGAGCA-----CTTCCACC 519
Db 363 nLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPheLeuLysSe 383
QY 520 GTGCAGG---CCTGCAGGTGGAGTTTCATCAACCCCATCTTCGAGTTCTCGCGGCCCATGC 576
Db 383 rLeuArgLysProPheGly-AppPheMetGluProLysPheGluPheAlaValLysPheA 403
QY 577 GCGCGCTGGCGCTGGACGACGCTGTAGTACGCCCTGTCTCATCGCCATCAACATCTTCTCGG 636
Db 403 snAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIleLeuSerG 423
QY 637 CGACCGCGCCCACTGCAGGAGCGCGCGCGCTGGAGCGCTGCAGACGCCCTACGCTGG 696
Db 423 lYAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAsnLeuLeuG 443
QY 697 AGGCGCTGCTGCTCTACACGCGCATCAAGAGCGCGCAG 734
Db 443 lnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGlu 455
```

RESULT 2

```
US-10-821-234-1096
; Sequence 1096, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1096
; LENGTH: 1467
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1096
```

```
Alignment Scores:
Pred. No.: 0.000293 Length: 1467
Score: 140.50 Matches: 81
Percent Similarity: 36.16% Conservative: 17
Best Local Similarity: 29.89% Mismatches: 101
Query Match: 10.23% Indels: 74
DB: 1 Gaps: 16
```

US-10-712-629b-2 (1-746) x US-10-821-234-1096 (1-1467)

```
QY 49 CTTGTGGGCGCGGACGAGCGAGCTCAGCGCTCTGGGCTGGGGCTTCCCTGGTGGA 108
Db 285 ProAlaGlyProLysGlyGluProGlySer-----ProGlyGluAsnGlyAlaPro 301
```

```
QY 109 TCTGAGGCGAGGAG-----CCAGGCTCCGGGAAGCGAGGGTGTCCAGCTAAC 158
Db 302 GlyGlnMetGlyProArgGlyLeuProGlyGluArgGlyArgProGlyAlaProGlyPro 321
QY 159 AGCGCTCAAGAACTAATGATCCACGAGTTGGTGGC-----GGCCCAACTGC- 205
Db 322 AlaGlyAlaArgGlyAsnAspGlyAlaThrGlyAlaAlaGlyProProGlyProThrGly 341
QY 206 -----AGTGCAACAAACGCTCTCTTCTCCGACC 232
Db 342 ProAlaGlyProProGlyPheProGlyAlaValAlaGlyAlaLysGlyGluAlaGlyProGln 361
QY 233 ACCCAAACTACGCCCTGGCCCTGGGGG-----CAGACCCCACTGTCAGATGCC 286
Db 362 GlyProArgGlySerGluGlyProGlnGlyValArgGlyGluProGlyProGlyPro 381
QY 287 GCCAGCAAGCTTCCGCTTCCGAGCTGCGGAGCTGCGCATCATCTCAGTCCAGGAGATCGTGG 346
Db 382 AlaGlyAlaAlaGlyProAlaGlyAsnProGlyAlaAspGlyGln-ProGlyAlaLysG 401
QY 347 ACTTCGCTAAGCA-----AGTCCCTGGTTTCTCGCAGCT---GGGCGGG 388
Db 401 YAlaAsnGlyAlaProGlyIleAlaGlyAlaProGlyPheProGlyAlaArgGlyProSe 421
QY 389 AGGACCATGATCCGCT-----CCTGAAGGCATCCACTATCGAGATCATGCTGCTAGAGA 442
Db 421 rGlyProGlnGlyProGlyGlyProProGly--ProLysGlyAsnSer-----Gly 437
QY 443 CAGCCAGGCGCTTACACCAACGAGACAGAGTGTATCACCTTCTTGGAGACTTCACCTACAG 502
Db 438 GluProGlyAlaProGlySerLysGlyAspThrGly-----Ala 450
QY 503 CAAGGACGACTTCCACCGCTGCAGGCGCTGCAGGTGGAGTTTCATCAACCCCATCTTTCAGTT 562
Db 451 LysGlyGluProGlyProVal-----GlyValGlnGlyPro----- 462
QY 563 CTCGCGGCGCATCGCGCGCTGGGCT-----GGAGGAGCTGAGTACGCGCCT 610
Db 463 -----ProGlyProAlaGlyGluGluGlyLysArgGlyAlaArgGly 476
QY 611 GCTCATCGCATCAACATCTTCTCGCGCGACGCGCCCAACGTCGAGAGCGCGGCGCGCT 670
Db 477 GluProGlyProThrGlyLeuProGlyProProGlyGluArgGlyGlyProGlySerArg 496
QY 671 GGA-----GGCGTTGCAGCAGCGCTACGTGGAGGCGCTGCTGTCTCTACAC 715
Db 497 GlyPheProGlyAlaAspGlyValAlaGlyPro---LysGlyProAla----- 511
QY 716 GGCATCAAGAGCGCGCAGGACGCT 742
Db 512 GlyGluArgGlySerProGlyProAla 520
```

RESULT 3

```
US-11-015-546A-20
; Sequence 20, Application US/11015546A
; Publication No. US20050250126A1
; GENERAL INFORMATION:
; APPLICANT: GAO, ZEREN
; APPLICANT: SHEPARD, PAUL O.
; APPLICANT: FOX, BRIAN A.
; APPLICANT: HOLLOWAY, JAMES L.
; APPLICANT: JASPERS, STEPHEN R.
; APPLICANT: APPELBY, MARK
; TITLE OF INVENTION: Ztnf13, A TUMOR NECROSIS FACTOR
; FILE REFERENCE: 03-24
; CURRENT APPLICATION NUMBER: US/11/015,546A
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: 60/530,185
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 614
```


Db 490 ---GlnGlyAlaAlaSerArgProAlaIleLeuTyrAlaLeuLeuSerSerLeuLys 508
QY 270 CCAGTCCCGAGATGCCCGCAGCAACGCTTTGGCCCACTTCACGGA----- 314
Db 509 AlaValProArgProArgSerArgCysLeuCysArgGlnHisArgProValGlnLeuCys 528
QY 315 GGTGGCCATCATCTAGTCCAGGATCGTGGAC-----TTCGC 353
Db 529 AlaProHisArg-ThrCysArgGluAlaLeuAspValLeuAlaIleThrValAlaPheLe 548
QY 354 TAAGCAAGTGCCTGTTCTCTGAGCTGGCCGGGAGGACCATCGCTCTCTGAAG-- 411
Db 548 uArgAenLeuProSerPheTrpGlnLeuProGlnAspGlnArgLeuLeuGlnG 568
QY 412 -----SerSerSerGlyGly-----GCATCCACTATCGAGATC-- 429
Db 568 yCysTrpGlyProLeuPheLeuLeuGlyLeuAlaGlnAspAlaValThrPheGluValAl 588
QY 430 -----ATGCTGTAGAGACAGCCAGCGCGCTACAA 458
Db 588 aGluAlaProValProSerIleLeuLysIleLeuLeuGlu-GluPro----- 604
QY 459 CCAGGAGACAGATGTATCACCTTCTTGAGGACTTCACCTACGACGAGGACGACTTCCA- 517
Db 605 -----SerSerSerGlyGly-----SerGlyGlnLeuProA 615
QY 518 -----CGTGCAGCGCTGCAGGTGGAGTTCATCAACCCCATCTTCGAGTTCTCG 566
Db 615 spArgProGlnProSerLeuAlaAla-ValGlnTrpLeuGlnCysCysLeuGlu----- 632
QY 567 CGGGCCATCGGGCGGTGGCCCTGGACGAGCTGAGTACGCCCTGCTCATCGCCATCAAC 626
Db 633 ---SerPheTrpSerLeuGluLeuSerProLysGluTyrAlaCysLeuLysGlyThrIle 651
QY 627 ATCTTCTCGGCGACCGGCCCAACGTGACGAGCGCGCGCGCTGGAGCGTTGACGAG 686
Db 652 LeuPheAsnProAspValProGlyLeuGlnAlaAlaSerHisIleGlyHisLeuGlnGln 671
QY 687 ---CCCTACGTGGAGCGCTGC-----TGTCCTACACGCGCATCAAG 725
Db 672 GluAlaHis-TrpValLeuCysGluValLeuGluProTrpCysPro-----AlaAlaG 689
QY 726 AGCGCCGAGGACC 738
Db 689 nGlyArgLeuThr 693

RESULT 7
US-10-821-234-964
; Sequence 964, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Pre-eclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 964
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-964

Alignment Scores:
Pred. No.: 0.0378 Length: 1166
Score: 112.00 Matches: 46
Percent Similarity: 35.12% Conservative: 13

Best Local Similarity: 27.38% Mismatches: 49
Query Match: 8.15% Indels: 60
DB: 1 Gaps: 7
US-10-712-629b-2 (1-746) x US-10-821-234-964 (1-1166)
QY 43 CAGTCCACTGTGTGGGCGGCGAGGCGCAGCAGCTCAGCTCTGGGCTTGGGCTTCCCTC 102
Db 105 GlnGlyProGlnGlyProLysGlyAsp-----ProGlyProGlyIlePro 120
QY 103 GGTGATCTGAGCAGCAGCAGCGGCTCCGGGGAAGGCGAGGTGTCCAGCTAAC----- 158
Db 121 GlyArgAsn--Gly---AspProGlyIleProGlyGlnProGlySerProGlySerProG 139
QY 159 -----AGCGGCTCAAGAACTAATGATCCA----- 182
Db 139 lyProProGlyIleCysGluSerCysProThrGlyProGlnAsnTyrSerProGlnTyrA 159
QY 183 -----GCAGTTTGGTGGC-----GGCCCAA 201
Db 159 spSerTyrAspValLysSerGlyValAlaValGlyGlyLeuAlaGlyTyrProGlyProA 179
QY 202 CTGACGTGCAACAAACGCTCTTCTCCGACACGCCCAAGATCACGCCCTTGGCCCTGGGC 261
Db 179 lagly----- 180
QY 262 GCAGACCCCGGCTCCGAGATGCCCGCCAGCAACGCTTGGCCCACTTCACGAGCTGGCC 321
Db 181 -----ProProGlyProProGlyProGlyProGlyThr-----SerGlyH 193
QY 322 ATCATCTCAGTCCAGGATCGTGACTTCGGCTTAAGCAAGTGCCTGCTTCTTCGACAGCTG 381
Db 193 isProGlySerProGlySerProGlyTyrGlnGlyProGlyGluProGlyGlnAlaG 213
QY 382 GCGCGGAGAGACAGATCCCTCTGTAAGGATCCACTATCGAGATCATGCTGCTAGAG 441
Db 213 lyProSerGlyProProGlyProProGlyAlaIleGlyProSerGlyProAlaGlyLysA 233
QY 442 ACAGCAGCGCGCTACAAACA 461
Db 233 spGlyGluSerGlyArgPro 239

RESULT 8
US-11-053-100-48
; Sequence 48, Application US/11053100
; Publication No. US20050255554A1
; GENERAL INFORMATION:
; APPLICANT: CHILKOTI, Ashutosh
; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION
; FILE REFERENCE: 4176-101 CIP
; CURRENT APPLICATION NUMBER: US/11/053,100
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US 09/812,382
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,659
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 48
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(624)
; OTHER INFORMATION: pET15b-SD3-ELP1-60-throm-Batrogen receptor ligand binding domain
US-11-053-100-48

Alignment Scores:
Pred. No.: 0.0455 Length: 624
Score: 111.00 Matches: 39

Percent Similarity:	40.97%	Conservative:	20
Best Local Similarity:	27.08%	Mismatches:	61
Query Match:	8.08%	Indels:	24
DB:		Gaps:	4
US-10-712-629B-2 (1-746) x US-11-053-100-49 (1-624)			
QY	52 GTGGGCGCAGGCGCAGCAGCAGCTCAGCCTCTGGG---CCTGGGCTTCC---CCTGGT 105		
DB	284 ValGlyValProGlyValGlyValProGlyValProGlyValProGlyValProGly 303		
QY	106 GGATCTGAGCAGCAGCAGCAGGCTCCGGGAAGCGAGGT-----147		
DB	304 GlyGlyValProGlyTrpProSerSerGlyGlyGlySerIleGlyProLeuVal 323		
QY	148 -----GTCCAGCTAACAGCGCTCAAGAACTAATGATCCAGCAG 186		
DB	324 ProArgGlySerHisMetSerLysLysAsnSerLeuAlaLeuSerLeuThrAlaAspGln 343		
QY	187 TTGGTGGCGGCCCAACTGCAGTGCACAAACGCTCTCTCCGACCCAGCCCAAGTCACG 246		
DB	344 MetValSerAlaLeuLeuAspAlaGluProProIleLeuTySerGlu-----359		
QY	247 CCTGGCCCTGGCGCAGACCCCGAGTCCCGAGATGCCCGCAGCAACGCTTTGCCAC 306		
DB	360 -----TyrAspProThrArgProPheSerGluAlaSerMetMetGlyLeu 374		
QY	307 TTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGTAAGCAAGTGCCT 366		
DB	375 LeuThrAsnLeuAlaAspArgGluLeuValHisMetIleAsnTrpAlaLysArgValPro 394		
QY	367 GGTTCCTCAGCTGGCGCGGAGGACCATCGCCCTCTCGAAGCATCCACTATCGAG 426		
DB	395 GlyPheValAspLeuThrLeuHisAspGlnValHisLeuLeuGluCysAlaTrpLeuGlu 414		
QY	427 ATCATGCTGCTA 438		
DB	415 IleLeuMetIle 418		
RESULT 9			
US-11-053-100-49			
; Sequence 49, Application US/11053100			
; Publication No. US2005025554A1			
GENERAL INFORMATION:			
; APPLICANT: CHILKOTI, Ashutosh			
; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION			
; FILE REFERENCE: 4176-101 CIP			
; CURRENT APPLICATION NUMBER: US/11/053,100			
; CURRENT FILING DATE: 2005-02-08			
; PRIOR APPLICATION NUMBER: US 09/812,382			
; PRIOR FILING DATE: 2001-03-20			
; PRIOR APPLICATION NUMBER: US 60/190,659			
; PRIOR FILING DATE: 2000-03-20			
; NUMBER OF SEQ ID NOS: 58			
; SOFTWARE: PatentIn version 3.3			
; SEQ ID NO 49			
; LENGTH: 774			
; TYPE: PRT			
; ORGANISM: Artificial			
FEATURE:			
OTHER INFORMATION: Synthetic Construct			
FEATURE:			
NAME/KEY: MISC FEATURE			
LOCATION: (1)..(774)			
OTHER INFORMATION: pET15b-SD5-ELPI-90-throm-Estrogen receptor ligand binding domain			
US-11-053-100-49			
Alignment Scores:			
Pred. No.:	0.0452	Length:	774
Score:	111.00	Matches:	39
Percent Similarity:	40.97%	Conservative:	20
Best Local Similarity:	27.08%	Mismatches:	61
Query Match:	8.08%	Indels:	24
US-10-712-629B-2 (1-746) x US-11-059-982-1 (1-1560)			
QY	43 CAGTCACTGTGGCGCGCAGGCGCAGCAGCTCAGCCTCTGGGCT-----90		
DB			
RESULT 10			
US-11-059-982-1			
; Sequence 1, Application US/11059982			
; Publication No. US20050255507A1			
GENERAL INFORMATION:			
; APPLICANT: Jenkins, Robert B.			
; APPLICANT: Yang, Ping			
; APPLICANT: Thibodeau, Steve			
; APPLICANT: Wang, Liang			
; APPLICANT: Schaid, Daniel			
; TITLE OF INVENTION: CYTOGENETICALLY DETERMINED DIAGNOSIS AND			
; TITLE OF INVENTION: PROGNOSIS OF PROLIFERATIVE DISORDERS			
; FILE REFERENCE: 07039-505001			
; CURRENT APPLICATION NUMBER: US/11/059,982			
; CURRENT FILING DATE: 2005-02-17			
; PRIOR APPLICATION NUMBER: US 60/545,573			
; PRIOR FILING DATE: 2004-02-17			
; NUMBER OF SEQ ID NOS: 50			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 1			
; LENGTH: 1560			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-11-059-982-1			
Alignment Scores:			
Pred. No.:	0.0446	Length:	1560
Score:	111.00	Matches:	79
Percent Similarity:	35.22%	Conservative:	27
Best Local Similarity:	26.25%	Mismatches:	94
Query Match:	8.08%	Indels:	101
DB:		Gaps:	18

Db 256 GlnValProVal-----SerGlyTyrLeuAlaSerAlaAlaGlyProSerGluProVal 273
 QY 91 -----GGGCTTCCCT----- 102
 Db 274 ThrLeuAlaSerAlaGlyValSerProGlnGlyAlaGlyLeuValleGlnLysLeu 293
 QY 103 -----GGTGATCTGAGGAGCC 120
 Db 294 SerAlaAlaValAlaThrThrLeuAenGlyAenSerValPheGlyGlyAlaAla 313
 QY 121 AGCCAGGCTCCGGGAGGAGGCTGTCAGCTAACAGCGCTCAAGAACTAATGATC 180
 Db 314 SerAlaProThrGlyThrProSerGlyGlnProLeuAlaValAlaProGlyLeuGlySer 333
 QY 181 CAGCAGTTGGTG---GCGGCCCAACTGCAGTGCACAAACCGTCTCTCCAGCAGCC 237
 Db 334 SerProLeuValProAla-ProAenValleuHisArgThrProThrProleGlnPr 353
 QY 238 AAGTCAAGC-----CTGGGCCCTGGGGCGGAGACCCCGAGTCCGAGATGCC 285
 Db 353 olysProAlaGlyValLeuProProLysLeuTyrGlnLeuThrProLysProPheAlaPr 373
 QY 286 CCGCAGCAACGCTTTCGCCACTTCAGGAGCTGGCATCATCTCAGTCAGGAGATCGTG 345
 Db 373 oAlaGlyAlaThrLeu---ThrileGlnGlyGluProGlyAlaLeuProGlnGlnProly 392
 QY 346 GACTTCGCTAAGCAAGTGCCTGTTCTTCAGCTGGGGCGGAGG-----AC 393
 Db 392 sAlaProGlnAenLeuThrPheMetAlaAlaGlyLysAlaGlyGlnAenValleuSe 412
 QY 394 CAGATCGCCCTCTGAAGCATCTACTATCGAGATCGTCTGAGACAGCCAGCGCC 453
 Db 412 rGlyPheProAla-----ProAlaLeuGlnAlaAenValPheLysGlnProAl 429
 QY 454 TACACACAGCAGACAGATGATACCTCTTCAGGACTTCACCTACAGCAAGGAGACT 513
 Db 429 aThrThr-----GlyAlaAl 435
 QY 514 TCCACCGTCAGGCTCGAGGTGGAGTTCATCAACC-----CATCTTCGAGTTCT 564
 Db 435 aProPro---GlnProProGlyAlaLeuSerLysProMetSerValHisLeu----- 451
 QY 565 CGCGGGCCATCGCGCGCTGGCGCTGGAGCAGCTGAGTACGCCCTGCTCATCGCCATCA 624
 Db 452 -----LeuAenGlnGlySerSerileValleProAla-----Gl 463
 QY 625 ACATCTTCTC-----GGCCGACCGGCC----- 646
 Db 463 nHisMetLeuProGlyGlnAenGlnPheLeuLeuProGlyAlaProAlaValGlnLeuPr 483
 QY 647 -CAACGTGCGAGGCGCGGCGCGT-----GGAGGC-----GTTGCGAGCAGC 687
 Db 483 oGlnGlnLeuSerAlaLeuProAlaAenValGlyGlyGlnlleLeuAlaAlaAlaPr 503
 QY 688 CTAAGTGGAGCGCTGCTGCTCTACAC-----GCGCATCAAGAGCGCGCAGGACC 738
 Db 503 oHisThrGlyGlyGlnLeuAlaAenProleLeuThrAenGlnAenLeuAlaGlyPr 523
 QY 739 A 739
 Db 523 o 523

RESULT 11
 US-10-821-234-1182
 ; Sequence 1182, Application US/10821234
 ; Publication No. US20050255114A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Labat, Ivan
 ; APPLICANT: Stache-Crain, Birgit
 ; APPLICANT: Andarmani, Susan
 ; APPLICANT: Tang, Y. Tom
 ; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
 ; FILE REFERENCE: 821A

; CURRENT APPLICATION NUMBER: US/10/821,234
 ; CURRENT FILING DATE: 2004-04-07
 ; PRIOR APPLICATION NUMBER: US 60/462,047
 ; PRIOR FILING DATE: 2003-04-07
 ; NUMBER OF SEQ ID NOS: 1704
 ; SOFTWARE: PC_SEQ_genes Version 1.0
 ; SEQ ID NO 1182
 ; LENGTH: 1874
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-821-234-1182
 Alignment Scores:
 Pred. No.: 0.074 Length: 1874
 Score: 108.00 Matches: 67
 Percent Similarity: 33.92% Conservative: 10
 Best Local Similarity: 29.52% Mismatches: 82
 Query Match: 7.86% Indels: 70
 DB: 1 Gaps: 12
 US-10-712-629B-2 (1-746) x US-10-821-234-1182 (1-1874)
 QY 49 CTGTGGGGCCGACGAGGAGCAGCTCA-----GCCTCTGGGCTGGGCTTCCCT 102
 Db 1289 ProProGlyProArgGlyProSerGlyAlaProGlyAlaAspGlyProGlnGlyProPro 1308
 QY 103 GGTGGA-----TCTGAGCAGCAG- 122
 Db 1309 GlyGlylleGlyAenProGlyAlaValGlyGluLysGlyProGlyGluAlaGlyGlu 1328
 QY 123 CCAGGGCTCCGGGAGGAGGCTGTCAGCTTAACAGCGGCTCAAGAACTAATGATCCA 182
 Db 1329 ProGlyProSerGlyArgSerGlyProProGlyProLysGlyGluArgGlyGluLysGly 1348
 QY 183 GCAGTTGGTGGGGCCCAACTCAGTGCACAAACGCTCTCTCTCGACGAGCCCAAGT 242
 Db 1349 Glu-----SerGlyProSerGly----- 1354
 QY 243 CAGCGCTGGCCCTGGGGCGGAGACCCAGTCCCGAGATGCCCGCCAGCAGCAGCTTTCG 302
 Db 1355 -----AlaAlaGly-----ProProGlyProLysGlyProProGlyAspAspGly 1369
 QY 303 CCACCTTACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAAGT 362
 Db 1370 ProLysGlySerProGlyProValGlyPheProGlyAenProGly----- 1384
 QY 363 GCCTGGTTCTTCGAGCT-----GGGCGGGAGGAGCAGATCGCCCTCTGAAGCATC 416
 Db 1385 -----ProProGlyGluProGlyProAlaGlyGlnAspGlyProPro----- 1398
 QY 417 CACTATCAGATCATGCTGCTAG-AGACAGCCAGCGGCTACAAACAGCAGAGAGAGTGA 475
 Db 1399 -----GlyAspLysGlyAspAspGlyGluProGly-----GlnThrGly 1411
 QY 476 TCACCTTCTTGAAGGACTTCACCTACAGCAAGGAGCAGCTTCCACCGTGCAGGCTGCAGGT 535
 Db 1412 SerProGly-----ProThrGlyGluProGlyProSerGlyProProGly 1426
 QY 536 GGAGTTTCATCAACCCCATCTTCGAGTTCTCGGGGCCATCGCGCGGCTGGGCT---GGA 592
 Db 1427 LysArgGlyProPro-----GlyProAlaGlyProGluGly 1438
 QY 593 CGACGCTGAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGCGGAGCCGCGCCCAACGT 652
 Db 1439 ArgGlnGlyGluLysGlyAlaLysGlyGluAlaGlyLeuGluGlyProProGlyLysThr 1458
 QY 653 GCAGGAGCGCGCGCGCTGGA 673
 Db 1459 GlyProLleGlyProGlnGly 1465
 RESULT 12
 US-11-053-100-50
 ; Sequence 50, Application US/11053100

```
; Publication No. US2005025554A1
; GENERAL INFORMATION:
; APPLICANT: CHILKOTI, Ashutosh
; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION
; FILE REFERENCE: 4176-101 CIP
; CURRENT APPLICATION NUMBER: US/11/053,100
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US 09/812,382
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,659
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 51
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(775)
; OTHER INFORMATION: pET15b-SDS-EUP1-180-throm-Estrogen receptor ligand binding domain
US-11-053-100-50
Alignment Scores:
Pred. No.: 0.0813 Length: 1225
Score: 107.50 Matches: 32
Percent Similarity: 42.97% Conservative: 23
Best Local Similarity: 25.00% Mismatches: 60
Query Match: 7.82% Indels: 13
DB: 7 Gaps: 1
US-10-712-629b-2 (1-746) x US-11-053-100-50 (1-1225)
QY 55 GGGCCCGCAGGCGAGCAGCTCAGCTCTGGCCCTGGGGCTTCCCTGGTGGATCTGAG 114
Db 905 GlyValProGlyTrpProSerSerGlyLeuValProArgGlySerProGlyIleSerGly 924
QY 115 GCAGGCGAGCAGGGCTCCGGGAGGCGAGGGTGTCCAGCTACAGCGGCTCAAGAACTA 174
Db 925 GlyGlyGlyHisMetSerLysLysAsnSerLeuAlaLeuSerLeuThrAlaAspGln 944
QY 175 ATGATCCAGCAGTGTGGCGGCCCAACTGCAGTGCAACAAACGCTCTTCTCCGACCAG 234
Db 945 MetValSerAlaLeuAlaAspAlaGluProProIleLeuTy:SerGluTy: 961
QY 235 CCCAAGTCAGCCCTGGCCCTCGGGCGGAGAGCCCGAGTCCCGAGATGCCCGCCAGCAA 294
Db 962 -----AspProThrArgProPheSerGluAlaSer 971
QY 295 CGCTTTGCCACTTCACGGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCT 354
Db 972 MetMetGlyLeuLeuThrAsnLeuAlaAspArgGlyLeuValHisMetIleAsnTrpAla 991
QY 355 AAGCAAGTCGCTGTTTCCTGACAGCTGGCGCGGAGGACAGATCGCCCTCTCTGAAGGCA 414
Db 992 LysArgValProGlyPheValAspLeuThrLeuHisAspGlnValHisLeuLeuGluCys 1011
QY 415 TCCACTATCGAGATCATGTGCTA 438
Db 1012 AlaTrpLeuGluIleLeuMetIle 1019
RESULT 13
US-11-053-100-51
; Sequence 51, Application US/11053100
; Publication No. US2005025554A1
; GENERAL INFORMATION:
; APPLICANT: CHILKOTI, Ashutosh
; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION
; FILE REFERENCE: 4176-101 CIP
; CURRENT APPLICATION NUMBER: US/11/053,100
; CURRENT FILING DATE: 2005-02-08
; Publication No. US2005025554A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Brian A.
; APPLICANT: Holloway, James L.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: Znf14, A Tumor Necrosis Factor
; TITLE OF INVENTION: Receptor
; FILE REFERENCE: 03-17
; CURRENT APPLICATION NUMBER: US/10/967,527A
; CURRENT FILING DATE: 2004-10-18
; Publication No. US 09/812,382
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,659
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 51
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(775)
; OTHER INFORMATION: pET15b-SD6-EUP1-90-TEV-Estrogen receptor ligand binding domain
US-11-053-100-51
Alignment Scores:
Pred. No.: 0.115 Length: 775
Score: 105.50 Matches: 38
Percent Similarity: 41.38% Conservative: 22
Best Local Similarity: 26.21% Mismatches: 60
Query Match: 7.68% Indels: 25
DB: 7 Gaps: 5
US-10-712-629b-2 (1-746) x US-11-053-100-51 (1-775)
QY 52 GTGGGGCGCAGGCGAGCAGCAGCTCAGCTCTGGG---CCTGGGGCTTCC---CCTGGT 105
Db 434 ValGlyValProGlyValGlyValProGlyGlyGlyValProGlyValProGlyValProGly 453
QY 106 GGATCTGAGGCGAGCAGCGGCTCCGGGAA----- 138
Db 454 GlyGlyValProGlyTrpProSerSerGlyAspTy:AspIleProThrThrGluAsnLeu 473
QY 139 ---GGCAGGGTGTCCAGCTA-----ACAGCGGCTCAAGAACTAATGATCCAG 183
Db 474 TyrPheGlnGlyAlaHisMetSerLysLysAsnSerLeuAlaLeuSerLeuThrAlaAsp 493
QY 184 CAGTTGGTGGCGGCCCAACTGCAGTGCAACAAACGCTCTTCTCCGAGCAGCCCAAGTC 243
Db 494 GlnMetValSerAlaLeuLeuAspAlaGluProProIleLeuTy:SerGlu----- 510
QY 244 ACGCCCTGGCCCTGGCGCAGACCCCGAGTCCCGAGATGCCCGCAGCAACGCTTTGCC 303
Db 511 -----TyrAspProThrArgProPheSerGluAlaSerMetMetGly 524
QY 304 CACTTCAGGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAAGTG 363
Db 525 LeuLeuThrAsnLeuAlaAspArgGlyLeuValHisMetIleAsnTrpAlaValArgVal 544
QY 364 CTGTTTCTCAGCTGGCGCGGAGGAGACCATCGCCCTCTCTGAAGGCATCCACTATC 423
Db 545 ProGlyPheValAspLeuThrLeuHisAspGlnValHisLeuLeuGluCysAlaTrpLeu 564
QY 424 GAGATCATGCTGCTA 438
Db 565 GluIleLeuMetIle 569
RESULT 14
US-10-967-527A-30
; Sequence 30, Application US/10967527A
; Publication No. US20050256041A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Brian A.
; APPLICANT: Holloway, James L.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: Znf14, A Tumor Necrosis Factor
; TITLE OF INVENTION: Receptor
; FILE REFERENCE: 03-17
; CURRENT APPLICATION NUMBER: US/10/967,527A
; CURRENT FILING DATE: 2004-10-18
```


; PRIOR APPLICATION NUMBER: 60/511,698
; PRIOR FILING DATE: 2003-10-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 308
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-967-527A-30

Alignment Scores:
Pred. No.: 0.166 Length: 308
Score: 103.50 Matches: 46
Percent Similarity: 43.97% Conservative: 5
Best Local Similarity: 39.66% Mismatches: 37
Query Match: 7.53% Indels: 28
DB: 1 Gaps: 9

US-10-712-629B-2 (1-746) x US-10-967-527A-30 (1-308)

QY 110 CTGAGCAGCAGCCAGGCTCCGGGAAGCGAGGGTGTCCAGCTAACAGCGGCTCAAG 169
|||||
Db 205 LeuArgGlnAlaGlyAlaAlaProGlyGlnGly-----His-GlySerAr 219
QY 170 AACTAATGATCCAGCAGTTGGTGGCGGCCCACTGCAGTCGCAACAAACGCTCTCTCCG 229
|||||
Db 219 GCys-----LeuProGlyGlyGlyPro-----TyrGlnGlnCysLeuThrTrpAr 234
QY 230 ACCAGCCCAAGTCAGCCCTGCGCCCTGGCGGCACAGCCCGAGTC---CCGAGATGCC 286
|||||
Db 234 gProArgProArgHisAlaLeuGly---GlyArgAspProGluAlaAspProLeuAlaGl 253
QY 287 GCCAGCAACGCTTGGCCCACTTCAAGGAGCTGGSCCATCATCTCAGTCCAGGAGATCGTGG 346
|||||
Db 253 YThrSerThrLysCys-----TrpHisArgProAlaProGly---ThrVa 267
QY 347 ACTTCGCTAAGCAAGTCCTGGTTCTCTCGAGCT-----GGGCGC 385
|||||
Db 267 lLeuGlyThrAlaSerAlaLeuGlyProSerAlaSerGlnArgArgThrProLysGlyPr 287
QY 386 GGGAGGACCATGCGCCCTCTCTGAAGCATCCAC-----TATCGA 425
|||||
Db 287 oGlyProGluAlaProProHisHisSerHisLeuValTyrArg 302

RESULT 15

US-10-821-234-1431
; Sequence 1431, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1431
; LENGTH: 1366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1431

Alignment Scores:
Pred. No.: 0.16 Length: 1366
Score: 103.50 Matches: 79
Percent Similarity: 33.10% Conservative: 15
Best Local Similarity: 27.82% Mismatches: 97
Query Match: 7.48% Indels: 93

DB: 1 Gaps: 16
US-10-712-629B-2 (1-746) x US-10-821-234-1431 (1-1366)
QY 741 GCTGTCCTGCGGCCCTCTTTGATGCGCGTGTAGACAGCAGCGCTCCACACTAGGCTGCT 682
|||||
Db 585 AlaGlyProArgGlyGlu-----ArgGlyProProGlyGluSerGly 598
QY 681 GCAACGCTCCACGCGCGCGCTCTCTGCACGTTGGCGGTGCGCGGAGAAATGTTGA 622
|||||
Db 599 AlaAlaGlyProThrGlyProLeuGlySerArg-GlyProSerGlyProProGlyProAs 618
QY 621 TGGCGAT-----GAGCAGGGGCTACTCAGCGTCGTCAGGCCAGCCGCGCATGCCCC 568
|||||
Db 618 pGlyAsnLysGlyGluProGlyValGlyAlaValGlyThrAlaGlyProSerGlyPr 638
QY 567 GCGAGAACTC----- 558
Db 638 oSerGlyLeuProGlyGluArgGlyAlaAlaGlyIleProGlyGlyLysGlyGluLysGl 658
QY 557 -GAAGATGGGTTGATGAACTCCACCTGCAGGCTGCACGGTGAAGTCTGCTCTGCTGT 499
Db 658 yGluPro-GlyLeuArgGlyGluIleGlyAsnProGlyArg----- 671
QY 498 AGGTGAAGTCTCTCAAGAGGTGATACACTGTCTGCTGCTGTCTAGCGCTGCTGCTC 439
Db 672 -----AspGlyAlaArgGlyAlaHisGlyAlaValGlyA 683
QY 438 TAGCAGCATGATCTCGATAGTGATGCTTTCAGGA-----G 403
Db 683 laProGly-----ProAlaGlyAlaThrGlyAspArgGlyGluAlaG 697
QY 402 GCGCATCTGCTCTCCCGCCAGCTGCAGGAACCCAGGCACCTTGCTTACGGAAGTCCAC 343
Db 697 yAlaAlaGlyProAlaGlyProAla-----GlyProA 708
QY 342 GA---TCTCTGACTGAGATGATGCGCAGCTCCGTGAAGTGGGCAAGCGTTGCTGCGC 286
Db 708 rgGlySerProGlyGluArgGlyGluValGlyPro---AlaGlyProAsnGlyPheAlaG 727
QY 285 GGCATCTCGGACTGGGGGTCTGCGCCAGGGGCCAGGGCGTGACTTTGGGCT----- 233
Db 727 lProAlaGlyAlaAlaGly---GlnProGlyAlaLysGlyGluArgGlyAlaLysGlyP 746
QY 232 ---GGTCGGAGAGGAGCGCTTGTTCACCTGCAGTT----- 200
Db 746 roLysGlyGluAsnGlyValValGlyProThrGlyProValGlyAlaAlaGlyProAlaG 766
QY 199 -----GGCGCCGCCAACCACTGCTGGATCATTA-----GTTCTTGAGCCGCTGTAG 154
Db 766 lProAsnGlyProProGlyProAlaGlySerArgGlyAspGlyGlyProProGlyMet 786
QY 153 CTGGACACCTCGCCTT-----CCCCGAGCCCTGCTGCTGCTGCTCAG 109
Db 786 hrGlyPheProGlyAlaAlaGlyArgThrGlyProProGlyProSerGlyIle--SerGl 805
QY 108 TCCACAGGGGAAGCCCGCCAGGCCAGAGCTGAGCTGCTGCTGCTGCGGCCCGCCACAGG 49
Db 805 yProProGly---ProProGlyProAlaGlyLysGluGlyLeuArg---GlyProArgGl 823
QY 48 TGAC 45
Db 823 yAsp 824

Search completed: November 25, 2005, 22:32:04
Job time : 19.5 secs

The Fugate (copy)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 22:03:31 ; Search time 165 Seconds
(without alignments)
1754.883 Million cell updates/sec

Title: US-10-712-629B-18
Perfect score: 3704
Sequence: 1 VTQCSCVQAAGEVGLTGH.....MDRAVFQAVKAAGALQEAQ 693

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pap.*
- 2: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pap.*
- 3: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pap.*
- 4: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pap.*
- 5: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pap.*
- 6: /cgn2_6/prodata1/pubpaa/US11_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3682	99.4	693	US-10-452-858C-93	Sequence 93, Appl
2	3682	99.4	1182	US-10-024-368-6	Sequence 6, Appl
3	3510	94.8	1207	US-10-024-368-5	Sequence 5, Appl
4	3095.5	83.6	1189	US-10-024-368-4	Sequence 4, Appl
5	3089.5	83.4	1189	US-10-122-013-17	Sequence 17, Appl
6	3088.5	83.4	984	US-10-024-368-2	Sequence 2, Appl
7	3077.5	83.1	1189	US-10-024-368-3	Sequence 3, Appl
8	729	19.7	1265	US-10-408-765A-795	Sequence 795, App
9	728.5	19.7	2055	US-10-193-874-20	Sequence 20, Appl
10	728.5	19.7	2321	US-10-193-874-17	Sequence 17, Appl
11	728.5	19.7	2358	US-10-193-874-14	Sequence 14, Appl
12	728.5	19.7	2540	US-10-193-874-16	Sequence 16, Appl
13	728.5	19.7	2552	US-10-193-874-15	Sequence 15, Appl
14	719	19.4	2210	US-10-719-993-589	Sequence 589, App
15	714	19.3	2353	US-10-719-993-588	Sequence 588, App
16	714	19.3	2535	US-10-719-993-587	Sequence 587, App
17	700.5	18.9	1417	US-10-408-765A-1992	Sequence 1992, Ap
18	606	16.4	671	US-10-264-049-2730	Sequence 2730, Ap
19	603.5	16.3	1212	US-10-170-385-19	Sequence 19, Appl
20	550.5	14.9	854	US-11-097-143-17979	Sequence 17979, A
21	539	14.6	1193	US-10-193-874-18	Sequence 18, Appl
22	448.5	12.1	952	US-10-755-889-640	Sequence 640, App
23	313.5	8.5	872	US-10-437-963-119587	Sequence 119587,
24	297	8.0	747	US-10-425-114-38701	Sequence 38701, A
25	297	8.0	941	US-10-424-599-210074	Sequence 210074,
26	280.5	7.6	874	US-10-437-963-158893	Sequence 158893,
27	279.5	7.5	363	US-10-425-114-42461	Sequence 42461, A

ALIGNMENTS

RESULT 1

US-10-452-858C-93
; Sequence 93, Application US/10452858C
; Publication No. US20040086945A1
; GENERAL INFORMATION:

; APPLICANT: Sreekrishna, Kotikanyadanam
; APPLICANT: Gerwe, Gina S.
; TITLE OF INVENTION: HAIRLESS PROTEIN-INTERACTING PARTNER COMPLEXES AND METHODS THERE
; FILE REFERENCE: 8956P
; CURRENT APPLICATION NUMBER: US/10/452,858C
; CURRENT FILING DATE: 2003-06-02
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 93
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-452-858C-93

Query Match		99.4%;	Score 3682;	DB 4;	Length 693;
Best Local Similarity		99.4%;	Pred. No. 1.5e-278;		
Matches 689;		Conservative 1;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1	VTQCSCVQAAGEVGLTGH	SQKRRRSPLEEKQLEEDSSATSEGGGPGPEASLNKGL	60	
Db	1	VTQCSCVQAAGEVGLTGH	SQKRRRSPLEEKQLEEDSSATSEGGGPGPEASLNKGL	60	
QY	61	AKHLLSGLDRLCRLLRKEREALAWAQREGQGPANTEDSPGIPHCCSRCHHGLFNTHWC	120		
Db	61	AKHLLSGLDRLCRLLRKEREALAWAQREGQGPANTEDSPGIPHCCSRCHHGLFNTHWC	120		
QY	121	SHCSHRLCVACGRAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAEIS	180		
Db	121	SHCSHRLCVACGRAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAEIS	180		
QY	181	VTMHQWAKFDIRGHCFQVDARVWAPGDGGQKEPTEKTPPTPQSCNGDSNRTKDIXE	240		
Db	181	VTMHQWAKFDIRGHCFQVDARVWAPGDGGQKEPTEKTPPTPQSCNGDSNRTKDIXE	240		
QY	241	ETPDSTESPAEDGAGRSPLPCSLCELLASTAVKLCGHDRTHMAFAPVTPALPSDDRT	300		
Db	241	ETPDSTESPAEDGAGRSPLPCSLCELLASTAVKLCGHDRTHMAFAPVTPALPSDDRT	300		
QY	301	NILDSIIIAQVVERKIQEALGRLRAGSLRKLGLSLPLSPVTRLSPPGALLWLOQPRPK	360		
Db	301	NILDSIIIAQVVERKIQEALGRLRAGSLRKLGLSLPLSPVTRLSPPGALLWLOQPRPK	360		
QY	361	HGFHLFQBHWRCQPVLSVGIQKTLRLSLWGMALGTGGQVQSITALGPPQPTNLDDSTA	420		
Db	361	HGFHLFQBHWRCQPVLSVGIQKTLRLSLWGMALGTGGQVQSITALGPPQPTNLDDSTA	420		

Db 361 HGPHLFOEHWROQGPVLVSGIQKTLRLSLWGMALGTLGGQVQTLTALGPPQPTNLDSTA 420
QY 421 FWEGFHPETRPKLDGSGVLLHRLTGLDKDASRVQNLVSSLPPEYCAHQGKLNLAASYLP 480
Db 421 FWEGFHPETRPKLDGSGVLLHRLTGLDKDASRVQNLVSSLPPEYCAHQGKLNLAASYLP 480
QY 481 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 540
Db 481 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 540
QY 541 GLDGEGLWSPGSGTSTVMHVFRADAQRIIRRFQMVCPAGAGTLEPGAGSCYLDAGLR 600
Db 541 GLDGEGLWSPGSGTSTVMHVFRADAQRIIRRFQMVCPAGAGTLEPGAGSCYLDAGLR 600
QY 601 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACLCHOG 660
Db 601 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACLCHOG 660
QY 661 ASLPPDHRMLYAQMRAVFOAVKAAVGALEAK 693
Db 661 ASLPPDHRMLYAQMRAVFOAVKAAVGALEAK 693

RESULT 2

US-10-024-368-6
; Sequence 6, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1182
; TYPE: PRT
; ORGANISM: Mouse
US-10-024-368-6

Query Match 99.4%; Score 3682; DB 4; Length 1182;
Best Local Similarity 99.4%; Pred. No. 3, 2e-278;
Matches 689; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VTQCSCVQAAGEVGLTGHSSQKRRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 60
Db 490 VTQCSCVQAAGEVGLTGHSSQKRRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 549
QY 61 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGI PHCCSRCHHGLFNTHWC 120
Db 550 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGI PHCCSRCHHGLFNTHWC 609
QY 121 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 180
Db 610 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 669
QY 181 TVMHQAWAKFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPPQSCNGDSNRTKDIKE 240
Db 670 TVMHQAWAKFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPPQSCNGDSNRTKDIKE 729
QY 241 ETPDSTESPAEDGAGRSPLCPSCCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
Db 730 ETPDSTESPAEDGAGRSPLCPSCCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 789
QY 301 NILDSIIAQQVVERKIQEALGPGIRAGSLRKGSLPLSPVTRTSLSPGALLMLQEPKPK 360
Db 790 NILDSIIAQQVVERKIQEALGPGIRAGSLRKGSLPLSPVTRTSLSPGALLMLQEPKPK 849

QY 361 HGPHLFOEHWROQGPVLVSGIQKTLRLSLWGMALGTLGGQVQSLTALGPPQPTNLDSTA 420
Db 850 HGPHLFOEHWROQGPVLVSGIQKTLRLSLWGMALGTLGGQVQTLTALGPPQPTNLDSTA 909
QY 421 FWEGFHPETRPKLDGSGVLLHRLTGLDKDASRVQNLVSSLPPEYCAHQGKLNLAASYLP 480
Db 910 FWEGFHPETRPKLDGSGVLLHRLTGLDKDASRVQNLVSSLPPEYCAHQGKLNLAASYLP 969
QY 481 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 540
Db 970 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 1029
QY 541 GLDGEGLWSPGSGTSTVMHVFRADAQRIIRRFQMVCPAGAGTLEPGAGSCYLDAGLR 600
Db 1030 GLDGEGLWSPGSGTSTVMHVFRADAQRIIRRFQMVCPAGAGTLEPGAGSCYLDAGLR 1089
QY 601 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACLCHOG 660
Db 1090 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACLCHOG 1149
QY 661 ASLPPDHRMLYAQMRAVFOAVKAAVGALEAK 693
Db 1150 ASLPPDHRMLYAQMRAVFOAVKAAVGALEAK 1182

RESULT 3

US-10-024-368-5
; Sequence 5, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: Rat
US-10-024-368-5

Query Match 94.8%; Score 3510; DB 4; Length 1207;
Best Local Similarity 95.1%; Pred. No. 9e-265;
Matches 659; Conservative 11; Mismatches 21; Indels 2; Gaps 2;

QY 1 VTQCSCVQAAGEVGLTGHSSQKRRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 60
Db 517 VTQCSCVQAAGEVGLTGHSSQKRRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 574
QY 61 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGI PHCCSRCHHGLFNTHWC 120
Db 575 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGI PHCCSRCHHGLFNTHWC 634
QY 121 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 180
Db 635 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 694
QY 181 TVMHQAWAKFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPPQSCNGDSNRTKDIKE 240
Db 695 TVMHQAWAKFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPPQSCNGDSNRTKDIKE 754
QY 241 ETPDSTESPAEDGAGRSPLCPSCCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
Db 755 ETPDSTESPAEDGAGRSPLCPSCCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 814
QY 301 NILDSIIAQQVVERKIQEALGPGIRAGSLRKGSLPLSPVTRTSLSPGALLMLQEPKPK 360

Db 815 NILDSTIAQVVERKIQEKALGPGLRAGSLRKLGLSLPLSPVTRLSPPGALLWLQEPRPK 874
QY 361 HGPHLFOEHWROQOPVLVSGIQKTILRLSLWGMALGTLGGQVQSLTALGPQPTNLDSTA 420
Db 875 HGPHLFOEHWROQOPVLVSGIQKTILRLSLWGMALGTLGGQVQSLTALGPQPTNLDSTA 934
QY 421 FMEGFSHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 480
Db 935 FMEGFSHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 994
QY 481 LGTLHPLPQLWAAVGVNSHRHGLTKNLCEVSDLSILVHAELPPEYCAHOGKLNLSYLP 540
Db 995 LGTLHPLPQLWAAVGVNSHRHGLTKNLCEVSDLSILVHAELPPEYCAHOGKLNLSYLP 1054
QY 541 GLDGEGLWSPGSGTSTVHVFRQAQRIIRRFLOMVCPCAGAGTLEPGAGSCVLDAGLR 600
Db 1055 GLDGEGLWSPGSGTSTVHVFRQAQRIIRRFLOMVCPCAGAGTLEPGAGSCVLDAGLR 1114
QY 601 RLREEWGVCWTLLQAPGEAVLVPAGAPHQVQGLVSTISVTHFLSPETSALSACLCHQ 660
Db 1115 RLREEWGVCWTLLQAPGEAVLVPAGAPHQVQGLVSTISVTHFLSPETSALSACLCHQ 1174
QY 661 ASLPPDHRMLYAQMDDRAVFAVKAAYGALQEA 693
Db 1175 ASLPPDHRMLYAQMDDRAVFAVKAAYGALQEA 1207

RESULT 4

US-10-024-368-4
; Sequence 4, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-368-4

Query Match 83.6%; Score 3095.5; DB 4; Length 1189;
Best Local Similarity 83.2%; Pred. No. 2.1e-232;
Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGVGLTGHSSQKRRSPL-EEKQLEEDSSATSEGGGPGPEASLNKG 59
Db 494 LAQCSCQAAGEGGCHACHSQVRRSPGLGELQEEEDTATNSSSEEGPGSPDRLSTG 553
QY 60 LAKHLLSGLDRLCLRLKERELAWAQREGOGPANTEPSGPHCCSRCHGLFNTHWR 119
Db 554 LAKHLLSGLDRLCLRLKERELAWAQREGOGPANTEPSGPHCCSRCHGLFNTHWR 613
QY 120 CSHCSHRLCVACGRITAGAGNREKTSQEQHTDDCAQEAAGHAACSLILTFVSSQALAE 179
Db 614 CPCSRLHLCVACGRVAGTGRAREKAGFQESAECTQEAAGHAACSLMLTFVSSQALAE 673
QY 180 STMVQAWAKFDIRGHCFQVDARVWAPGDGGQKEPTKPTPTPQSCNGDSNRKTDIK 239
Db 674 STAMHQVWVKFDIRGHCFQVDARVWAPGDGGQKEPTKPTPTPQSCNGDTHRTKSIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 299
Db 734 EETPDSAETPAEDRAGRGLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 793

QY 300 TNILDSIIAQVVERKIQEKALGPGLRAGSLRKLGLSLPLSPVTRLSPPGALLWLQEPRP 358
Db 794 TNILDSIIAQVVERKIQEKALGPGLRAGSLRKLGLSLPLSPVTRLSPPGALLWLQEPRP 853
QY 359 -PKHGPHLFOEHWROQOPVLVSGIQKTILRLSLWGMALGTLGGQVQSLTALGPQPTNLD 417
Db 854 CPMRGPHLFOEHWROQOPVLVSGIQKTILRLSLWGMALGTLGGQVQSLTALGPQPTNLD 913
QY 418 STAFWEGFSHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLS 477
Db 914 STTFWEGFSHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLS 973
QY 478 YLPLGLTLHPLPQLWAAVGVNSHRHGLTKNLCEVSDLSILVHAELPPEYCAHOGKLN 537
Db 974 YLPLGLTLHPLPQLWAAVGVNSHRHGLTKNLCEVSDLSILVHAELPPEYCAHOGKLN 1033
QY 538 FLUGLDEGLWSPGSGTSTVHVFRQAQRIIRRFLOMVCPCAGAGTLEPGAGSCVLDAG 597
Db 1034 FLUGLDEGLWSPGSGTSTVHVFRQAQRIIRRFLOMVCPCAGAGTLEPGAGSCVLDAG 1093
QY 598 LRRRLREEWGVCWTLLQAPGEAVLVPAGAPHQVQGLVSTISVTHFLSPETSALSACL 657
Db 1094 LRRRLREEWGVCWTLLQAPGEAVLVPAGAPHQVQGLVSTISVTHFLSPETSALSACL 1153
QY 658 HQGASLPPDHRMLYAQMDDRAVFAVKAAYGALQEA 693
Db 1154 HQGASLPPDHRMLYAQMDDRAVFAVKAAYGALQEA 1189

RESULT 5

US-10-122-013-17
; Sequence 17, Application US/10122013
; Publication No. US20030077614A1
; GENERAL INFORMATION:
; APPLICANT: Christiano, Angela
; TITLE OF INVENTION: NUCLEIC ACIDS FOR INHIBITING HAIRLESS PROTEIN EXPRESSION AND MET
; FILE REFERENCE: USE THEREOF
; FILE REFERENCE: 0575/62637A
; CURRENT APPLICATION NUMBER: US/10/122,013
; CURRENT FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-122-013-17

Query Match 83.4%; Score 3089.5; DB 4; Length 1189;
Best Local Similarity 83.0%; Pred. No. 6.3e-232;
Matches 578; Conservative 40; Mismatches 75; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGVGLTGHSSQKRRSPL-EEKQLEEDSSATSEGGGPGPEASLNKG 59
Db 494 LAQCSCQAAGEGGCHACHSQVRRSPGLGELQEEEDTATNSSSEEGPGSPDRLSTG 553
QY 60 LAKHLLSGLDRLCLRLKERELAWAQREGOGPANTEPSGPHCCSRCHGLFNTHWR 119
Db 554 LAKHLLSGLDRLCLRLKERELAWAQREGOGPANTEPSGPHCCSRCHGLFNTHWR 613
QY 120 CSHCSHRLCVACGRITAGAGNREKTSQEQHTDDCAQEAAGHAACSLILTFVSSQALAE 179
Db 614 CPCSRLHLCVACGRVAGTGRAREKAGFQESAECTQEAAGHAACSLMLTFVSSQALAE 673
QY 180 STMVQAWAKFDIRGHCFQVDARVWAPGDGGQKEPTKPTPTPQSCNGDSNRKTDIK 239
Db 674 STAMHQVWVKFDIRGHCFQVDARVWAPGDGGQKEPTKPTPTPQSCNGDTHRTKSIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 299
Db 734 EETPDSAETPAEDRAGRGLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 793
QY 300 TNILDSIIAQVVERKIQEKALGPGLRAGSLRKLGLSLPLSPVTRLSPPGALLWLQEPRP 358

Db 794 TNLDSIIAQVVERKIQEALGPGLRAGPGLRKGGLPLSPVPRPLPPPGALLWLQEPQP 853
QY 359 -PKHGPHLQFQHWROQOPVLVSGIQKTLRLSLMGMEALGTGQVQSLTALGPPQPTNLD 417
Db 854 CPRGRPHLQFQHWROQOPVLVSGIQKTLQGNLWGTALGALGGQVQALSPLGPPQSSLG 913
QY 418 STAFWEGFSGHPETRPKLDDEGSVLLHRLTGDKDASRVQNLVSSLPLPEYCAHOGKLNLAS 477
Db 914 STTFWEGFSGHPETRPKLDDEGSVLLHRLTGDKDASRVQNLVSSLPLPEYCAHOGKLNLAS 973
QY 478 YLPLGLTLPLPQLWAAVGNVSHRGHLGTKNLCVEVDLSILVHAERQALPQWYRAQKD 537
Db 974 YLPPGLALPLPQLWAAVGNVSHRGHLGTKNLCVEVDLSILVHAERQALPQWYRAQKD 1033
QY 538 FLSGLDGEGWSPGSGTSTVWHVFRADQRIIRRFQWVCPAGAGTLEPGAPGSCYLDAG 597
Db 1034 FLSGLDGEGWSPGSGTSTVWHVFRADQRIIRRFQWVCPAGAGTLEPGAPGSCYLDAG 1093
QY 598 LRRRLREEVGSCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQALC 657
Db 1094 LRRRLREEVGSCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQALC 1153
QY 658 HQGASLPDPHRLMYAQMDRAVFOAVKAAVGALEQAK 693
Db 1154 HOGPSLPDPCHLLY AQMDWAVFOAVKVAVGTLOEAK 1189

RESULT 6

US-10-024-368-2
; Sequence 2, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-368-2

Query Match 83.4%; Score 3088.5; DB 4; Length 984;
Best Local Similarity 83.0%; Pred. No. 5.8e-232;
Matches 578; Conservative 40; Mismatches 75; Indels 3; Gaps 2;

QY 1 VTQCSCVQAAGEVGLTGHQSRRSPL-EEKQLEEDSSATSEEGGGGPGPEASLNGK 59
Db 289 LAQCSCQAAGGGGCHACHSQVRRSPGLGELQOEEDTATNSSSEEGGPGPSRLSTG 348
QY 60 LAKHLISGLDRLCRLLRKEREALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWR 119
Db 349 LAKHLISGLDRLCRLLRKEREALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWR 408
QY 120 CSHCSHRLCVACGRIAGAGKNNREKTSQEOHTDDCAQEAAGAACSLTLTOFVSSQALAE 179
Db 409 CPRCSHRLCVACGRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTOFVSSQALAE 468
QY 180 STVHQAWAKFDTRGHCFCQVDARVAPGQGGQKTEPTTPTPQPSNCNGDSNRTKDIK 239
Db 469 STAMHQVWKFDIRGHCFCQVDARVAPGQGGQKTEPTTPTPQPSNCNGDSNRTKDIK 528
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDIRIHMAFAPVTPALPSDDR 299
Db 529 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDIRIHMAFAPVTPALPSDDR 588

QY 300 TNLDSIIAQVVERKIQEALGPGLRAGSRLRKGSLPLSPVTRTLSPGALLWLQEPR- 358
Db 589 TNLDSIIAQVVERKIQEALGPGLRAGSRLRKGSLPLSPVTRTLSPGALLWLQEPQP 648
QY 359 -PKHGPHLQFQHWROQOPVLVSGIQKTLRLSLMGMEALGTGQVQSLTALGPPQPTNLD 417
Db 649 CPRGRPHLQFQHWROQOPVLVSGIQKTLQGNLWGTALGALGGQVQALSPLGPPQSSLG 708
QY 418 STAFWEGFSGHPETRPKLDDEGSVLLHRLTGDKDASRVQNLVSSLPLPEYCAHOGKLNLAS 477
Db 709 STTFWEGFSGHPETRPKLDDEGSVLLHRLTGDKDASRVQNLVSSLPLPEYCAHOGKLNLAS 768
QY 478 YLPLGLTLPLPQLWAAVGNVSHRGHLGTKNLCVEVDLSILVHAERQALPQWYRAQKD 537
Db 769 YLPPGLALPLPQLWAAVGNVSHRGHLGTKNLCVEVDLSILVHAERQALPQWYRAQKD 828
QY 538 FLSGLDGEGWSPGSGTSTVWHVFRADQRIIRRFQWVCPAGAGTLEPGAPGSCYLDAG 597
Db 829 FLSGLDGEGWSPGSGTSTVWHVFRADQRIIRRFQWVCPAGAGTLEPGAPGSCYLDAG 888
QY 598 LRRRLREEVGSCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQALC 657
Db 889 LRRRLREEVGSCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQALC 948
QY 658 HQGASLPDPHRLMYAQMDRAVFOAVKAAVGALEQAK 693
Db 949 HOGPSLPDPCHLLY AQMDWAVFOAVKVAVGTLOEAK 984

RESULT 7

US-10-024-368-3
; Sequence 3, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-04-07
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-368-3

Query Match 83.1%; Score 3077.5; DB 4; Length 1189;
Best Local Similarity 82.9%; Pred. No. 5.4e-231;
Matches 577; Conservative 37; Mismatches 79; Indels 3; Gaps 2;

QY 1 VTQCSCVQAAGEVGLTGHQSRRSPL-EEKQLEEDSSATSEEGGGGPGPEASLNGK 59
Db 494 LAQCSCQAAGGGGCHACHSQVRRSPGLGELQOEEDTATNSSSEEGGPGPSRLSTG 553
QY 60 LAKHLISGLDRLCRLLRKEREALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWR 119
Db 554 LAKHLISGLDRLCRLLRKEREALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWR 613
QY 120 CSHCSHRLCVACGRIAGAGKNNREKTSQEOHTDDCAQEAAGAACSLTLTOFVSSQALAE 179
Db 614 CPRCSHRLCVACGRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTOFVSSQALAE 673
QY 180 STVHQAWAKFDTRGHCFCQVDARVAPGQGGQKTEPTTPTPQPSNCNGDSNRTKDIK 239
Db 674 STAMHQVWKFDIRGHCFCQVDARVAPGQGGQKTEPTTPTPQPSNCNGDSNRTKDIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDIRIHMAFAPVTPALPSDDR 299

Db 734 EETPDSAEPAEDRAGRPFCPSCELLASTAVKLCGLHDIRHMAPVTPALPSDDRI 793
Qy 300 TNLDSIIAOWVERKIOEKALGCLGRLAGSLPLSPVTRLSPPCALLWLOEPR- 358
Db 794 TNLDSIIAOWVERKIOEKALGCLGRLAGSLPLSPVTRLSPPCALLWLOEPP 853
Qy 359 PKHGPHLFOEHRWQGPVLVSGIQKTLRLSLMGMEALGTGGVQSLSLTALGPPQPTNLD 417
Db 854 CPRRGPHLFOEHRWQGPVLVSGIQKTLRLSLMGMEALGTGGVQSLSLTALGPPQPTNLD 913
Qy 418 STAFWEGFSPHETPKLDEGSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKUNLAS 477
Db 914 STTFWEGFSPHETPKLDEGSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKUNLAS 973
Qy 478 YLPLGLTLHLEPQLMAAYGVNSHRLGHLGKNCVSVDSILSVHAEALPPEYRAQKD 537
Db 974 YLPPGLALRLEPQLMAAYGVNSHRLGHLGKNCVSVDSILSVHAEALPPEYRAQKD 1033
Qy 538 FLSLDGEGLWSPGSGTSTVHVFRQAQRIIRFLQWCPAGAGTLEPGAPGSCYLDAG 597
Db 1034 FLSLDGEGLWSPGSGTSTVHVFRQAQRIIRFLQWCPAGAGTLEPGAPGSCYLDAG 1093
Qy 598 LRRLREEVGSCWTLLOAPGEAVLPAGAPHQVQGLVSTVTOHFLSPETSALSALC 657
Db 1094 LRRLREEVGSCWTLLOAPGEAVLPAGAPHQVQGLVSTVTOHFLSPETSALSALC 1153
Qy 658 HQGASLPDPHMLYAQMDRAVFAVKAACALQ 693
Db 1154 HQGASLPDPHMLYAQMDRAVFAVKAACALQ 1189

RESULT 8

US-10-408-765A-795
; Sequence 795, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 795
; LENGTH: 1265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-795

Query Match 19.7%; Score 729; DB 4; Length 1265;
Best Local Similarity 29.2%; Pred. No. 1.9e-47;
Matches 219; Conservative 117; Mismatches 270; Indels 144; Gaps 25;
Qy 61 AKHLLSGLDRLCRLLREREALAWAQREGQGPAMTDESPGPHCCSRCHGLFNTHWRC 120
Db 528 SKYILDIIGDKFCQLVTSEKTSLSWVKDAK-IWKRAVRGVREMDCAEATLFNTHWVC 586
Qy 121 SHCSHRLCVACGRAGAKNREKTSQEQHT-DDCAQEAAGHAACSILITQFVSSQALAE 179
Db 587 QKGFVVCCLDYK----AKERKSSRDKELYAMKVKQGPQPHDKHLMPTQIIPGSLTDL 642
Qy 180 STVMHQAQAFDIRGHCF-----QVDARVWAPGDGQQ-----KEPTE 218
Db 643 LDAMHTLRKYGIKSHCHTCKNQNLQVGNFPTMNGSVQVLQNLVNLHNSKISLCMPESQQ 702

Qy 219 KTPPTQPSCNGSDNRTKDIKE---TPDSTESP----- 249
Db 703 NTFPKSEK--NGSSPESDVGTNDKLTTPESQPLHLADLAQKAREEKKENKELTLEN 760
Qy 250 -----AEDGAGR-SPLPC-----PSCELLASTAVKLCGLHDIRHMAPV- 289
Db 761 QIKEREQDNSESPNGRTSPLVSNQNEQSTLRLDLLTTTAGKLRVSGTDAIAPAPVYSM 820
Qy 290 -TPALSDDDRITNLDIIIAOVVERKI-----QEKALGPGURAGSGLRKGL 334
Db 821 GAPSSKSGTMTNLDIIIASVVVENKI PPSKTSKINVKPELKEEPEESIISAVENNKLYS 880
Qy 335 SLPLSVTRLSPPGALLMLQEPKPHGPHLFOEHRWQGPVLVSGIQKTLRLSLMGMEA 394
Db 881 DIPHSWICEK----HILMLKDYKNSSNMKLKFECKKQGPAPVSVGHKQKNTLSLWKAES 935
Qy 395 LG-TLGGVQSLSLTALGPPQPTNLDSTAFWEGFSPHETPKLDEGSVLLHRTLGDKD- 450
Db 936 ISLDFGDHQAADLLNCKDSIISANVKEFWDFGFEVSKRQKSGETVVLKL-----KDWPS 991
Qy 451 -----ASRVQNLVSSLPPEYCAHOGKUNLASVLPGLTLHP-LPEPOLWAAYGVNSH 501
Db 992 GEDFTKMPARVEDLLKSLPLPEYCNPEGKFNASHLP-GFFVRPDLGPRCLCSAYGVAA 1050
Qy 502 RGH-LGTKNLCEVSDLSILVHAEALPPEYRAQKDFLSGLDGEGLWS----- 549
Db 1051 KOHDIGTTLNHFVSDVNLVYVGIAGKNGILSKAGILKKPEEDLDDILRKLKDSSE 1110
Qy 550 -PGSQTSTVHVFRQAQRIIRFLQWVC-PAGAGTL---EPGAPGSCYLDAGLRRLRE 604
Db 1111 IPGA-----LWHIYAGKVDKIREFLQKISKEQGLEVLPEHDPIRDQSWYVYVKKLRQLLE 1166
Qy 605 EWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTVTOHFLSPETSALSALCQHGASLP 664
Db 1167 EYGVRTCTLIQFLGDAIVLPAGALHQVQNFHSCIQVTEDFVSPEHLVESPHLTQELRLK 1226
Qy 665 P-----DHRMLYAQMDRAVFAVKAACALQ 690
Db 1227 BEINYDDKL--QVKNILYHAVKEMVRLK 1253

RESULT 9

US-10-193-874-20
; Sequence 20, Application US/10193874
; Publication No. US20030064396A1
; GENERAL INFORMATION:
; APPLICANT: Jenapharm GmbH & Co. KG
; TITLE OF INVENTION: Comodulators of Nuclear Receptors
; FILE REFERENCE: 52145
; CURRENT APPLICATION NUMBER: US/10/193,874
; CURRENT FILING DATE: 1998-11-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 2055
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-193-874-20

Query Match 19.7%; Score 728.5; DB 4; Length 2055;
Best Local Similarity 29.1%; Pred. No. 4e-47;
Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;
Qy 61 AKHLLSGLDRLCRLLREREALAWAQREGQGPAMTDESPGPHCCSRCHGLFNTHWRC 120
Db 1317 SKYILDIIGDKFCQLVTSEKTSLSWVKDAK-IWKRAVRGVREMDCAEATLFNTHWVC 1375
Qy 121 SHCSHRLCVACGRAGAKNREKTSQEQHT-DDCAQEAAGHAACSILITQFVSSQALAE 179
Db 1376 QKGFVVCCLDYK----AKERKSSRDKELYAMKVKQGPQPHDKHLMPTQIIPGSLTDL 1431
Qy 180 STVMHQAQAFDIRGHCF-----QVDARVWAPGDGQQ-----KEPTE 218

Db 1432 LDAMHTLRBYGKSHCHCTNKQNLQVGNFPTMNGVSQVLQNVLNHNSKISLCMPESQOQ 1491
 QY 219 KTPPTPQPCNGSDNRKDIKEE---TPDSTESP----- 249
 Db 1492 NTPPKSEK--NGGSSPESDVGTDNKLTPPESQPLHLADLAQKAREEKENKELTLEN 1549
 QY 250 -----AEDGAGR-SPLPC-----PSCELLASTAVKLCIGHDRHMAFAPV--- 289
 Db 1550 QIKEEREQDNSESPNGRTSPLVSNQNEQSTLRDLTTTAGKLKRVGSTDAGIAFAFVYSM 1609
 QY 290 -TPALPSDDRIITNLDISIQAQVVERKI-----QBKALGPGLRAGSGLR 331
 Db 1610 GAPSSKSGRTWPNILDDIIASVVENKIIPSKTSKINVKPELKEEPEESIISAVDENNKLY 1669
 QY 332 KGLSLPLSPVRTRLSPPGALLWLQEPKPHGPHLFQEHWRQGPVLVSGIQKTLRLSLMG 391
 Db 1670 S--DIPHSWICEK-----HILWLKDYKNSSNNWKLFECKWKQGPAPVVGHVHKKOMNISLWK 1722
 QY 392 MEALG-TLGGQVQSLTALGPQPTNLDSTAFWEGFSPHETPRPKLDGSGVLLHRTLGDKD 450
 Db 1723 AESISLDFGDHQAADLLNCKDSIISNANVKEFDGFEVSKRQNKSGETVVLKL-----KD 1778
 QY 451 -----ASRVQNLVSSLPPEYCAHQKLNLASYLPLGLTLHP--LEPOLWAAYGV 498
 Db 1779 WPSGEDFTKMPARYEDLLKSLELPEYCNPEGKFNLAHLUP-GFFVPRDLGPRLC SAYGV 1837
 QY 499 NSHRGH-LGTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLSGLDGEGLMS----- 549
 Db 1838 VAAKHIDGTTNLHIEVSDVNILVYVGIAGKNGILSKAGILKKFEEDLDDILKRLKD 1897
 QY 550 -----PSQSTSTVHVHFRADAQRIIRFLQWVC-PAGAGTL---EPGAPGSCYLDAGLRRR 601
 Db 1898 SSEIPGA-----LWHIYAGKDVDTIREFLQKISKEQGLEVLPEHDPIRDQSWYVYNNKKLQR 1953
 QY 602 LREEMGVSWTLLQAPCEAVLPAGAPHOVQGLVSTISVTQHFLSPETSALSQAOLCHOQA 661
 Db 1954 LLEBYGVTRCTLIQFLGDAIVLPAGALHQQVNFHSCIQVTEDFVSPHELVESFHLTQELR 2013
 QY 662 SLPPP-----DHRMLYAQMRAVFOAKAAVGAALQ 690
 Db 2014 LKKEEINYDKL---QVKNILYHAVKEMVRALK 2043

RESULT 10
 US-10-193-874-17
 ; Sequence 17, Application US/10193874
 ; Publication No. US20030064396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jenapharm GmbH & Co. KG
 ; TITLE OF INVENTION: Comodulators of Nuclear Receptors
 ; FILE REFERENCE: 52145
 ; CURRENT APPLICATION NUMBER: US/10/193,874
 ; CURRENT FILING DATE: 1998-11-18
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 17
 ; LENGTH: 2321
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-193-874-17

Query Match 19.7%; Score 728.5; DB 4; Length 2321;
 Best Local Similarity 29.1%; Pred. No. 4.7e-47;
 Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;

QY 61 AKHLLSGDLRCLRLKREKREALAWAQREGGQGPANTEDSPCIHCCSRCHHGLFNTHWRC 120
 Db 1583 SKYILDIGDKFCQLVTSKLTALSWKKDAK-IWKRAVRGVREMCACEATLNFNIHWVC 1641
 QY 121 SHCSHRLVCACGRIAGAKNREKTSQSOHT-DDCAQEAAGHAACSLILITQFVSSQALAE 179
 Db 1642 QKCGFVCLDCYK-----AKERKSRDKELYAMWCKVQKQPHDKHLMPTQIIPGSVLTDL 1697

QY 180 STVMQAWAKPDIRGHCFE---QVDARVWAPDGGQO-----KEPTE 218
 Db 1698 LDAMHTLRBYGKSHCHCTNKQNLQVGNFPTMNGVSQVLQNVLNHNSKISLCMPESQOQ 1757
 QY 219 KTPPTPQPCNGSDNRKDIKEE---TPDSTESP----- 249
 Db 1758 NTPPKSEK--NGGSSPESDVGTDNKLTPPESQPLHLADLAQKAREEKENKELTLEN 1815
 QY 250 -----AEDGAGR-SPLPC-----PSCELLASTAVKLCIGHDRHMAFAPV--- 289
 Db 1816 QIKEEREQDNSESPNGRTSPLVSNQNEQSTLRDLTTTAGKLKRVGSTDAGIAFAFVYSM 1875
 QY 290 -TPALPSDDRIITNLDISIQAQVVERKI-----QBKALGPGLRAGSGLR 331
 Db 1876 GAPSSKSGRTWPNILDDIIASVVENKIIPSKTSKINVKPELKEEPEESIISAVDENNKLY 1935
 QY 332 KGLSLPLSPVRTRLSPPGALLWLQEPKPHGPHLFQEHWRQGPVLVSGIQKTLRLSLMG 391
 Db 1936 S--DIPHSWICEK-----HILWLKDYKNSSNNWKLFECKWKQGPAPVVGHVHKKOMNISLWK 1988
 QY 392 MEALG-TLGGQVQSLTALGPQPTNLDSTAFWEGFSPHETPRPKLDGSGVLLHRTLGDKD 450
 Db 1989 AESISLDFGDHQAADLLNCKDSIISNANVKEFDGFEVSKRQNKSGETVVLKL-----KD 2044
 QY 451 -----ASRVQNLVSSLPPEYCAHQKLNLASYLPLGLTLHP--LEPOLWAAYGV 498
 Db 2045 WPSGEDFTKMPARYEDLLKSLELPEYCNPEGKFNLAHLUP-GFFVPRDLGPRLC SAYGV 2103
 QY 499 NSHRGH-LGTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLSGLDGEGLMS----- 549
 Db 2104 VAAKHIDGTTNLHIEVSDVNILVYVGIAGKNGILSKAGILKKFEEDLDDILKRLKD 2163
 QY 550 -----PSQSTSTVHVHFRADAQRIIRFLQWVC-PAGAGTL---EPGAPGSCYLDAGLRRR 601
 Db 2164 SSEIPGA-----LWHIYAGKDVDTIREFLQKISKEQGLEVLPEHDPIRDQSWYVYNNKKLQR 2219
 QY 602 LREEMGVSWTLLQAPCEAVLPAGAPHOVQGLVSTISVTQHFLSPETSALSQAOLCHOQA 661
 Db 2220 LLEBYGVTRCTLIQFLGDAIVLPAGALHQQVNFHSCIQVTEDFVSPHELVESFHLTQELR 2279
 QY 662 SLPPP-----DHRMLYAQMRAVFOAKAAVGAALQ 690
 Db 2280 LKKEEINYDKL---QVKNILYHAVKEMVRALK 2309

RESULT 11
 US-10-193-874-14
 ; Sequence 14, Application US/10193874
 ; Publication No. US20030064396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jenapharm GmbH & Co. KG
 ; TITLE OF INVENTION: Comodulators of Nuclear Receptors
 ; FILE REFERENCE: 52145
 ; CURRENT APPLICATION NUMBER: US/10/193,874
 ; CURRENT FILING DATE: 1998-11-18
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 2358
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-193-874-14

Query Match 19.7%; Score 728.5; DB 4; Length 2358;
 Best Local Similarity 29.1%; Pred. No. 4.8e-47;
 Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;

QY 61 AKHLLSGDLRCLRLKREKREALAWAQREGGQGPANTEDSPCIHCCSRCHHGLFNTHWRC 120
 Db 1620 SKYILDIGDKFCQLVTSKLTALSWKKDAK-IWKRAVRGVREMCACEATLNFNIHWVC 1678
 QY 121 SHCSHRLVCACGRIAGAKNREKTSQSOHT-DDCAQEAAGHAACSLILITQFVSSQALAE 179

Db 1679 QKQFVVCLDYK-----AKERSSRDKELYAMKCVKGQPHDKHLMPTQIIPGSLTDL 1734
QY 180 STMHQAWAKFDIRGHCF-----QVDARVWAPDGGQ-----KEPTE 218
Db 1735 LDAMHTLREKYGIKSHCHCTNKQNLQVGNPPTMNGVSQVLQNLVHNSKISLCMPESQOQ 1794
QY 219 KPTPTPOPCNGSDNRTKIKER---TPDSTESP----- 249
Db 1795 NTPPKSEK---NGSSPESDVGTDNKLTTPESQSLHDLARQKAREEKENKELTLEN 1852
QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCIGHDRHMAFAPV--- 289
Db 1853 QIKEREQDNSESPNGRTSPLVSQNNQOGLTLDLTTTAGKLRVGSTDAGIAFAPVYSM 1912
QY 290 -TPALPSDDRIITNLDIIIAQVVERKI-----QBKALGPGLRAGSGLR 331
Db 1913 GAPSSKSGRTMPNILDIIASVVENKIPPSKTSKINVKPELKEEPESIIISAVDENNKLY 1972
QY 332 KGLSLPLSPVTRTLSPGALLMQEPKPHGPHLFOEHWROQOPVLVSGIQKTLRLSLWG 391
Db 1973 S--DIPHSWICEK-----HILMLKDYKNSNNWKLKFCWKQGOQPAVVSGVHKOMNISLWK 2025
QY 392 MEALG-TLGGQVQSLTALGPPQPTNLDSTAFWEGFSHPETRPKLDGSGVLLHRTLGDKD 450
Db 2026 ABESISLDFGDHQAADLLNCKDSIISANVKEFMDGFEVSKRQKNKSGETVWLK-----KD 2081
QY 451 -----ASRVQNLVSSLPPEYCAHQKLNLSYPLGLTLHP-LEPOLMAAYGV 498
Db 2082 WPSGEDFKTMTMPARYEDLLKSLPLPEYCNPEGKFNLAHLP-GFFVRPDLGPRLCRAYGV 2140
QY 499 NSHRGH-LGTKNLCVEVSDLSILVHAEQAQLPPWYRAQKDFLSGLDGEGLWS----- 549
Db 2141 VAAKHDI GTTNLHIEVSDVNNILVYVGIAGKNGILSKAGILKKPFEEDLDDILRKRLKD 2200
QY 550 -----PGSQSTVWHVFRADAQRIIRFLQWVC-PAGAGTL---EPGAPGSCYLDAGLRRR 601
Db 2201 SSEIFGA---LWHIYAGKDVKIREFLQKISKEQGLEVLPEHDPTRDQSWYVNNKLRQR 2256
QY 602 LREEMGVSWTLLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQAOLCHQGA 661
Db 2257 LLEEYGVRTCTLIQFLGDAIVLPAGALHQVQNFHSCIQVTEDFVSPHELVESPHLTQELR 2316
QY 662 SLPP-----DHRMLYAQMDBRAVFOAKAAGCALQ 690
Db 2317 LLKEEINYDDKL---QVKNILYHAVKEMVRALK 2346

RESULT 12

US-10-193-874-16
; Sequence 16, Application US/10193874
; Publication No. US20030064396A1
; GENERAL INFORMATION:
; APPLICANT: Jenapharm GmbH & Co. KG
; TITLE OF INVENTION: Comodulators of Nuclear Receptors
; FILE REFERENCE: 52145
; CURRENT APPLICATION NUMBER: US/10/193,874
; CURRENT FILING DATE: 1998-11-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-193-874-16

Query Match 19.7%; Score 728.5; DB 4; Length 2540;
Best Local Similarity 29.1%; Pred. No. 5.4e-47;
Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;
QY 61 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCCSRCHGLFNTHWRC 120
Db 1802 SKYILDIIIGDKFCQLVTSEKTSALSWVKDAK-IANKRAVRGVREMDACEATLFTNIHWVC 1860

QY 121 SHCSHRLCVACRIAGAGKNREKTSQEQHT-DDCAQEAGHAACSLILTOFVSSQAALAE 179
Db 1861 QKQFVVCLDYK-----AKERSSRDKELYAMKCVKGQPHDKHLMPTQIIPGSLTDL 1916
QY 180 STMHQAWAKFDIRGHCF-----QVDARVWAPDGGQ-----KEPTE 218
Db 1917 LDAMHTLREKYGIKSHCHCTNKQNLQVGNPPTMNGVSQVLQNLVHNSKISLCMPESQOQ 1976
QY 219 KPTPTPOPCNGSDNRTKIKER---TPDSTESP----- 249
Db 1977 NTPPKSEK---NGSSPESDVGTDNKLTTPESQSLHDLARQKAREEKENKELTLEN 2034
QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCIGHDRHMAFAPV--- 289
Db 2035 QIKEREQDNSESPNGRTSPLVSQNNQOGLTLDLTTTAGKLRVGSTDAGIAFAPVYSM 2094
QY 290 -TPALPSDDRIITNLDIIIAQVVERKI-----QBKALGPGLRAGSGLR 331
Db 2095 GAPSSKSGRTMPNILDIIASVVENKIPPSKTSKINVKPELKEEPESIIISAVDENNKLY 2154
QY 332 KGLSLPLSPVTRTLSPGALLMQEPKPHGPHLFOEHWROQOPVLVSGIQKTLRLSLWG 391
Db 2155 S--DIPHSWICEK-----HILMLKDYKNSNNWKLKFCWKQGOQPAVVSGVHKOMNISLWK 2207
QY 392 MEALG-TLGGQVQSLTALGPPQPTNLDSTAFWEGFSHPETRPKLDGSGVLLHRTLGDKD 450
Db 2208 ABESISLDFGDHQAADLLNCKDSIISANVKEFMDGFEVSKRQKNKSGETVWLK-----KD 2263
QY 451 -----ASRVQNLVSSLPPEYCAHQKLNLSYPLGLTLHP-LEPOLMAAYGV 498
Db 2264 WPSGEDFKTMTMPARYEDLLKSLPLPEYCNPEGKFNLAHLP-GFFVRPDLGPRLCRAYGV 2322
QY 499 NSHRGH-LGTKNLCVEVSDLSILVHAEQAQLPPWYRAQKDFLSGLDGEGLWS----- 549
Db 2323 VAAKHDI GTTNLHIEVSDVNNILVYVGIAGKNGILSKAGILKKPFEEDLDDILRKRLKD 2382
QY 550 -----PGSQSTVWHVFRADAQRIIRFLQWVC-PAGAGTL---EPGAPGSCYLDAGLRRR 601
Db 2383 SSEIFGA---LWHIYAGKDVKIREFLQKISKEQGLEVLPEHDPTRDQSWYVNNKLRQR 2438
QY 602 LREEMGVSWTLLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQAOLCHQGA 661
Db 2439 LLEEYGVRTCTLIQFLGDAIVLPAGALHQVQNFHSCIQVTEDFVSPHELVESPHLTQELR 2498
QY 662 SLPP-----DHRMLYAQMDBRAVFOAKAAGCALQ 690
Db 2499 LLKEEINYDDKL---QVKNILYHAVKEMVRALK 2528

RESULT 13

US-10-193-874-15
; Sequence 15, Application US/10193874
; Publication No. US20030064396A1
; GENERAL INFORMATION:
; APPLICANT: Jenapharm GmbH & Co. KG
; TITLE OF INVENTION: Comodulators of Nuclear Receptors
; FILE REFERENCE: 52145
; CURRENT APPLICATION NUMBER: US/10/193,874
; CURRENT FILING DATE: 1998-11-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 2552
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-193-874-15

Query Match 19.7%; Score 728.5; DB 4; Length 2552;
Best Local Similarity 29.1%; Pred. No. 5.4e-47;
Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;
QY 61 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCCSRCHGLFNTHWRC 120

Db 1814 SKYILDIIGDKFCQLVTSEKLTALSWVKDKAK-IANKRAVRGVREMCDAEATLNFHWWC 1872
 QY 121 SHCSHRLCVACGRIAGAGKREKTSQEOHT-DDCAQEAAGHAACSLILTFQVSSQALAE 179
 Db 1873 QKCGFVCLDCYK----AKERSSRDKELYAMKCVKGQPHDKHLMPTQIIPGSLVTL 1928
 QY 180 STVMHQAQAFDIRGHCF----QVDARVWAPGDGQO-----KEPTE 218
 Db 1929 LDAMHTLREKYGKSHCHCTKNQNLQVGNFPTMNGVSQVLQNLVHNSKISLCMPESQO 1988
 QY 219 KTPPTQPSCNGSDNKTKEB---TPDSTESP----- 249
 Db 1989 NTPPKSEK--NGGSSPESDVTGNKLTTPESQPLHLADLAQKAREEKENKELTLEN 2046
 QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCGLGHDRIHMAFAPV--- 289
 Db 2047 QIKEREQDNSESPNGRTSPLVSONNEQSTLRLDTTTAGKLRVSTDAIAFAPVYSM 2106
 QY 290 -TPALPSDDRTITNLSIIAQVVERKI-----QBKALGPGLRAGSLR 331
 Db 2107 GAPSSKSGRTMPNILDIIASVVENKIPPSKTKINVKPELKEEPEBSIIISAVDENNKLY 2166
 QY 332 KGLSLPLSPVTRTLSPGALLWQEPKPGFHLFOEHWQOGQPVLSGIGKTLRLSLWG 391
 Db 2167 S--DIPHSWICEK-----HILWLKDYKNSNNWKLFECKWQGPVAVVGVHKKNNISLWK 2219
 QY 392 MEALG-TLGGQVQSLTALGPPQPTNLDSTAFWEGFHPETRPKLDGSLVLLHRTLTGDKD 450
 Db 2220 AESISLDFGDHODLLNCKDSIISANVKEFWDGFEVSKRQKSGETVVKL---KD 2275
 QY 451 -----ASRVQNLVSSLPPEYCAHQKLNLSYLPGLTLHP-LEPOLMAAYGV 498
 Db 2276 WPSGEDFKTMWPARYEDLLKSLPEYCNPEGFNLASHLP-GFFVRPDLGRLCSAYGV 2334
 QY 499 NSHRGH-LGTKNLCVEVSDLSILVHAEALPPWYRAQKDFLSGLDGEGLWS----- 549
 Db 2335 VAAKDHIDGTNTNLHIEVSDVNILVYVGIAGNGILSKAGILKKFEEEDLDDILRKRLKD 2394
 QY 550 ----PGSOTSTVWHVFRQAQIRRFLOWVC-PAGAGTL---EPGAPGSCYLDAGLRRR 601
 Db 2395 SSIFGA---LWHIYAGKVDKIREFLQKISKEQGLEVLPEHDPTRDQSWYVYNNKLQR 2450
 QY 602 LREWGVSCWTLQAQAEVLVAPAGAPHQVGLVSTISVTHFLSPETSALSALQCHQGA 661
 Db 2451 LLEEYGVRTCTLIQFLGDAIVLAPAGALHQVNFHSCIQVTEDFVSPHVLVEFHLTQELR 2510
 QY 662 SLPP----DHRMLYAQMDRAVQAVKAAGALQ 690
 Db 2511 LKKEEINYDDKL---QVKNILYHAKVEMVRALK 2540

RESULT 14
 US-10-719-993-589
 ; Sequence 589, Application US/10719993
 ; Publication No. US20040265849A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001496
 ; CURRENT APPLICATION NUMBER: US/10/719,993
 ; CURRENT FILING DATE: 2003-11-24
 ; NUMBER OF SEQ ID NOS: 55342
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 589
 ; LENGTH: 2210
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-719-993-589

Query Match 19.4%; Score 719; DB 5; Length 2210;
 Best Local Similarity 30.0%; Pred. No. 2.4e-46;
 Matches 207; Conservative 108; Mismatches 250; Indels 124; Gaps 23;

QY 61 AKHLSGLGDRLCRLLRKERREALAWAQREGQGPAMTEDSPGIPHCSCRCHGLFNTHWPC 120
 Db 1515 SKYILDIIGDKFCQLVTSEKLTALPRVKDKAK-IANKRAVRGVREMCDAEATLNFHWWC 1573
 QY 121 SHCSHRLCVACGRIAGAGKREKTSQEOHT-DDCAQEAAGHAACSLILTFQVSSQALAE 179
 Db 1574 QKCGFVCLDCYK----AKERSSRDKELYAMKCVKGQPHDKHLMPTQIIPGSLVTL 1629
 QY 180 STVMHQAQAFDIRGHCFQVDARVWA---PDGGQOQKEPTKPTTPOPSNGSDNRRTK 236
 Db 1630 LDAMHTLREKYGKSHCHCTKNQNLQVGNFPTMNGVSQSQOQNTTPPKSEK--NGSSPES 1687
 QY 237 DIKEB---TPDSTESP-----AEDGAGR- 256
 Db 1688 DVTGNKLTTPESQPLHLADLAQKAREEKENKELTLENOIKEREQDNSESPNGRT 1747
 QY 257 SPLPC-----PSLCELLASTAVKLCGLGHDRIHMAFAPV---TPALPSDDRTITNLSI 306
 Db 1748 SPLVSONNEQSTLRLDTTTAGKLRVSTDAIAFAPVYSMGAPSSKSGRTMPNILDII 1807
 QY 307 IAQVVERKI-----QEKALGPGLRAGSLRKGJSLPLSPVTRTLSPPG 349
 Db 1808 IASVVENKIPPSKTKINVKPELKEEPEBSIIISAVDENNKLYS--DIPHSWICEK----- 1860
 QY 350 ALLWQEPKPGFHLFOEHWQOGQPVLSGIGKTLRLSLWGMEALG-TLGGQVQSLTAL 408
 Db 1861 HILWLKDYKNSNNWKLFECKWQGPVAVVGVHKKNNISLWKAESISLDFGDHODLLNC 1920
 QY 409 GPPQPTNLDSTAFWEGFHPETRPKLDGSLVLLHRTLTGDKD-----ASRVQNL 457
 Db 1921 KOSIISANVKEFWDGFEVSKRQKSGETVVKL---KWPSEGEFKTMWPARYEDL 1976
 QY 458 VSSLPPEYCAHQKLNLSYLPGLTLHP-LEPOLMAAYGVNSHRGH-LGTKNLCVEVS 515
 Db 1977 LKSLPLPEYCNPEGFNLASHLP-GFFVRPDLGRLCSAYGVVAAKHIDGTTNLHIEVS 2035
 QY 516 DLISILVHAEALPPWYRAQKDFLSGLDGEGLWS-----PGSOTSTVWHVFR 563
 Db 2036 DVNINILVYVGIAGNGILSKAGILKKFEEEDLDDILRKLDKOSSEIPGA---LWHIYAG 2091
 QY 564 ODAQIRRRFLOWVC-PAGAGTL---EPGAPGSCYLDAGLRRRLREWGVSCWTLQAQGE 619
 Db 2092 KVDKIREFLQKISKEQGLEVLPEHDPTRDQSWYVYNNKLQRLLLEEYGVRTCTLIQFLGD 2151
 QY 620 AVLVPAGAPHQVGLVSTISVTHFLSPE 648
 Db 2152 AIVLPAGALHQVNFHSCIQVTEDFVSPE 2180

RESULT 15
 US-10-719-993-588
 ; Sequence 588, Application US/10719993
 ; Publication No. US20040265849A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001496
 ; CURRENT APPLICATION NUMBER: US/10/719,993
 ; CURRENT FILING DATE: 2003-11-24
 ; NUMBER OF SEQ ID NOS: 55342
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 588
 ; LENGTH: 2353
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-719-993-588

Query Match 19.3%; Score 714; DB 5; Length 2353;
 Best Local Similarity 29.1%; Pred. No. 6.6e-46;
 Matches 219; Conservative 117; Mismatches 263; Indels 154; Gaps 27;

```

QY 61 AKHLLSGLDRCLRLKEREALAWAQREGQGPAMTEDSPGIPHCCSRCHHGLFNTHWRC 120
Db 1620 SKYILDIIGFKQCVLTSEKTSALSWKQAK-IWKRAVRGVREMCDAEATLFINHWVC 1678
QY 121 SHCSHRLCVACGRIAGAGKREXTGQEOHT-DDCAQEAAGHAACSLILTQFVSSQALAE 179
Db 1679 QKCGFVCLDCYK-----AKERSSRDKELYAMWKVKQGPQPHDKHLMPTQI-----LTDL 1729
QY 180 STVMHQAQAFDIRGHCF-----QVDARVWAPGDGGQ-----KEPTE 218
Db 1730 LDAMHTLREKYGIKSHCHCTNKQNLQVGNFPTWNGVSVLQNVLNHSNKISLCMPESQQ 1789
QY 219 KTPPTQPCNGSDSNRTKDIKEE---TPDSTESP----- 249
Db 1790 NTPPKSEK--NGGSPESDVTGNKLTTPESQSPHLWADLAQKAREEKENKELTEN 1847
QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCIGHDRIHMAFVY--- 289
Db 1848 QIKEREQDNSESPNGRTSPLVSQNEQGSTLRDLTTTAGKLRVGSTDAGIAFAPVYSM 1907
QY 290 -TPALPSDDRIITNILDSIIAQVVERKI-----QEKALGPGLRAGSGLR 331
Db 1908 GAPSSKSGRTMPNILDIIASVVVENKIPPSKTSKINVKPELKEEPEESIISAVDENNKLY 1967
QY 332 KGLSLPLSPVTRTLSPPGALLWLOEPRPKHGFHLFQEHWRQGPVLVSGIQKTLRLSLWG 391
Db 1968 S--DIPHSWICEK-----HILWLKDYKNSSNWKLFKECKQGPQPAVVGHVHKKMNIWLK 2020
QY 392 MEALG-TLGGVQVSLTALGPQPTNLDSTAFWEGFSPHETPRPKLDEGSVLLHRTLGDKD 450
Db 2021 AESISLDFGDHQAADLLNCKDSIISNANVKFEWDFEVSQRQKNGSETVVLK-----KD 2076
QY 451 -----ASRVONLVSSLPPEYCAHOGKLNLASVPLGLTLHP-LEPOLWAAYGV 498
Db 2077 WPSGEDFKTMPPARYEDLLKSLPLPEYCNPEGKFNLAHLUP-GFFVRPDLGPRLC SAYV 2135
QY 499 NSHRGH-LGTQNLCEVSDLSILVHAEALQPPWYRAQKDFLSGLDGEGLWS----- 549
Db 2136 VAAKDHIDIGTTLNHLIEVSDVNILVYVGIAGNGILSKAGILKKEFEEDLDDILRRLKD 2195
QY 550 -----PGSQSTVWHVFRQAQAIRPLQWVC-PAGAGTL---EPGAPGSCYLDAGLRRR 601
Db 2196 SBEIPGA---LWHIYAGKDVDKIREFLQKISKEQGLEVLPEHDPTRDQSWYVYVKKLOR 2251
QY 602 LREEMGVSWCTLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACLCHQGA 661
Db 2252 LLEEYGVTRCTLLIOFLGDAIVLPAGALHQVQNFHSCIQVTEDFVSPHELVESPHLTQELR 2311
QY 662 SLPP-----DHRMLYAQMRAVFOAVKAAYGALQ 690
Db 2312 LLKEEINYDDKL---QVKNILYHAKEMVRALK 2341

```

Search completed: November 25, 2005, 22:17:09
 Job time : 171 secs

The Fog Bank (1870)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 22:03:15 ; Search time 44 Seconds
(without alignments)
1515.414 Million cell updates/sec

Title: US-10-712-629B-18
Perfect score: 3704
Sequence: 1 VTQCSCVQAAGEVGLTGH.....MDRAVFQAVKAAGALQEAQ 693
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3682	99.4	1182	2 I48378	hairless protein -
2	604.5	16.3	1214	2 S28499	probable finger pr
3	338	9.1	851	2 D86254	hypothetical prote
4	337	9.1	906	2 T01440	hypothetical prote
5	246	6.6	730	2 B85013	hypothetical prote
6	246	6.6	730	2 T10539	hypothetical prote
7	240.5	6.5	1629	2 T06461	DNA-binding protei
8	237	6.4	1641	2 T10955	early nodulin bind
9	232	6.3	1701	2 T43213	ENBPI protein - ba
10	205.5	5.5	728	2 T05151	hypothetical prote
11	181.5	4.9	950	2 F86222	hypothetical prote
12	146.5	4.0	2414	2 A54277	transcription adap
13	145.5	3.9	2440	2 S39162	transcription coac
14	145.5	3.9	2441	2 S39161	CREB-binding prote
15	127.5	3.4	1479	2 T42710	mannose receptor,
16	122.5	3.3	594	2 S33561	ref(2)P protein -
17	122	3.3	3938	2 T42761	Bassoon protein -
18	121	3.3	778	2 I38487	tastin - human
19	119	3.2	711	2 S05381	VGF8a protein proc
20	117.5	3.2	862	2 T46289	hypothetical prote
21	116	3.1	907	2 A24938	hypothetical T2 pr
22	115	3.1	761	2 S20458	pqqf protein - kle
23	114	3.1	442	2 JC4978	oxidative stress p
24	113.5	3.1	2509	2 G01880	fatty-acid synthas
25	112.5	3.0	728	2 S71467	diacylglycerol kin
26	112.5	3.0	739	2 A83015	primosomal protein
27	112.5	3.0	1616	2 I37183	gene APXL protein
28	111.5	3.0	3942	2 T42730	Bassoon protein -
29	111	3.0	915	2 S36327	clathrin assembly

30	111	3.0	1170	2 A57650	repair protein xpg
31	110.5	3.0	572	2 T37128	hypothetical prote
32	110	3.0	1072	2 A37127	microtubule-associ
33	110	3.0	4861	2 S71752	giant protein p619
34	109.5	3.0	1356	2 A45445	janusin precursor,
35	108.5	2.9	1113	2 T00271	hypothetical prote
36	108	2.9	1123	1 WMBEH7	UL37 protein - hum
37	107.5	2.9	1673	2 I50806	complement compone
38	107.5	2.9	5149	2 F83345	probable non-ribos
39	107	2.9	892	2 B46203	mating type A alph
40	107	2.9	2044	2 T13704	still life protein
41	107	2.9	2133	2 T42763	coagulation factor
42	106.5	2.9	353	2 B72468	hypothetical prote
43	106.5	2.9	427	2 JN0785	Carbon catabolite
44	106.5	2.9	896	2 S36326	clathrin assembly
45	106.5	2.9	1045	2 JC5795	CDEP protein - hum

ALIGNMENTS

RESULT 1

I48378
hairless protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I48378
R;Cachon-Gonzalez, M.B.; Fenner, S.; Coffin, J.M.; Moran, C.; Best, S.; Stoye, J.P.
Proc. Natl. Acad. Sci. U.S.A. 91, 7717-7721, 1994
A;Title: Structure and expression of the hairless gene of mice.
A;Reference number: I48378; MUID:94329587; PMID:8052649
A;Accession: I48378
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1182 <RES>
A;Cross-references: UNIPROT:O61645; UNIPARC:UPI0000028C24; EMBL:Z32675; NID:G531706; PI

Query Match 99.4%; Score 3682; DB 2; Length 1182;
Best Local Similarity 99.4%; Pred. No. 3.3e-257;
Matches 689; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY	1	VTQCSCVQAAGEVGLTGH	SQKRRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL	60
DB	490	VTQCSCVQAAGEVGLTGH	SQKRRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL	549
QY	61	AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCCSRCHHGLFNTHWRC	120	
DB	550	AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCCSRCHHGLFNTHWRC	609	
QY	121	SHCSHRLCVACGRIAGAGKNRKETSQEOHTDDCAQEAAGAACSLILTQFVSSQALAEALS	180	
DB	610	SHCSHRLCVACGRIAGAGKNRKETSQEOHTDDCAQEAAGAACSLILTQFVSSQALAEALS	669	
QY	181	TYMHQAWAKFDIRGHCFQVQDARVWAPGGQGGQKEPTPTPPQPSGNGDSNRTKDIKE	240	
DB	670	TYMHQAWAKFDIRGHCFQVQDARVWAPGGQGGQKEPTPTPPQPSGNGDSNRTKDIKE	729	
QY	241	ETPDTSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFVPTPALPSDDRIT	300	
DB	730	ETPDTSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFVPTPALPSDDRIT	789	
QY	301	NILDSIIAQVVRKIQEALGRLRAGSLRGLSLPLSPVTRLSPPGALLWLQEPKPK	360	
DB	790	NILDSIIAQVVRKIQEALGRLRAGSLRGLSLPLSPVTRLSPPGALLWLQEPKPK	849	
QY	361	HGFHLFQEHWRQGPVLVSGIQKTLRLSLWGMEALGTGGVQSLTALGPQPTNLIDSTA	420	
DB	850	HGFHLFQEHWRQGPVLVSGIQKTLRLSLWGMEALGTGGVQSLTALGPQPTNLIDSTA	909	
QY	421	FWEGFSHPETRPKLDGSGVLLHRLTGDKASRVQNLVSSLPPLPEYCAHQGLNLSYLP	480	
DB	910	FWEGFSHPETRPKLDGSGVLLHRLTGDKASRVQNLVSSLPPLPEYCAHQGLNLSYLP	969	

QY 481 LGTLHLPLPQWAAVGVNSHRHGLCTKNLCVEVSDLSILVHAEALPWPYRAQKDFLS 540
Db 970 LGTLHLPLPQWAAVGVNSHRHGLCTKNLCVEVSDLSILVHAEALPWPYRAQKDFLS 1029
QY 541 GLDGEGLWSPGSGTSTVHVHVFRAQDAQRIRRFLOWKVPAGAGTLEPGAGSCYLDAGLR 600
Db 1030 GLDGEGLWSPGSGTSTVHVHVFRAQDAQRIRRFLOWKVPAGAGTLEPGAGSCYLDAGLR 1089
QY 601 RLREMGVSCWTLLOAPGAVLVPAGAPHQVQGLVSTISVTHFLSPETSALSQAQCHQG 660
Db 1090 RLREMGVSCWTLLOAPGAVLVPAGAPHQVQGLVSTISVTHFLSPETSALSQAQCHQG 1149
QY 661 ASLPPDHRMLYAQMRAVFAVKAAGALQAEK 693
Db 1150 ASLPPDHRMLYAQMRAVFAVKAAGALQAEK 1182
RESULT 2
S28499
probable finger protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004
C:Accession: S28499
R;Hoog, C.; Schalling, M.; Grunder-Brundell, E.; Daneholt, B.
submitted to the EMBL Data Library, June 1991
A:Description: Analysis of a murine germ cell-specific transcript that encodes a putative
A:Reference number: S28499
A:Accession: S28499
A:Molecule type: mRNA
A:Residues: 1-1214 <HOO>
A:Cross-references: UNIPROT:Q63679; UNIPARC:UPI000013774D; EMBL:X59993; NID:g57503; PIDN
A:Experimental source: strain Sprague Dawley
C:Keywords: DNA binding; zinc finger
Query Match 16.3%; Score 604.5; DB 2; Length 1214;
Best Local Similarity 26.7%; Pred. No. 2.5e-35;
Matches 195; Conservative 114; Mismatches 289; Indels 131; Gaps 23;
QY 61 AKHLLSGLRCLRLKEREALWAQREGQGPAMTSDSPGIPHCSCRHGLFNTHWRC 120
Db 502 AKVILANIGDFQMWISEKAMSTIEPHRQ-VANKEAVKGVREMCVDCDTTIFNLHWVC 560
QY 121 SHCSHRLCVACGRIAGKAKREKTSQEQHT-DDCAQEAAGAACSLILTOFVSSQALAE 179
Db 561 PRCGFGVCVDCYRL--KRKNCQGAAYKTFWIRCVKSIHEPENLMPQTIIIPGKALYDV 618
QY 180 STVHQWAKFDIRGHCFC-----QVDARVWAPDGGQOKEPTKPTTPQPSGNGDSNR 234
Db 619 GDVHSVRKAWGKANCPCSNROPKLFSKPALKEDLKQASLSGK-----PSLGTWVQQ 672
QY 235 TKDKEETPDSTRESADGAGRSLPCPS-----LCELLASTAVKLCGLGHDRIHMAFA 287
Db 673 SSPVLEPAAVCGEAPSPASNVKPI-CPANTSPLNWLADLTSGNVNK-----ENKERKOLTM 727
QY 288 PV-----TPALP-----SDRITN- 301
Db 728 PILKNETKLPPLPPLNKSSTVLHTFNSTLTTPVNSNNGSLRNLNSSTGKTENGLKNT 787
QY 302 --TILDSIIAQVVERKIQEALGPGLRAGSL-RKGLSLPLSPVTRLSPP-----CALL 352
Db 788 PKLLDDIFASLVQNK-----TSSDLSKRPOGLTIKPSILGFTDTHYMLCDNRLL 836
QY 353 WLQEPKPKHGFHLFOEHWROGQPVLVSGIQKTLRLSLWGMALGTLGG--QVQSLTALGP 410
Db 837 CLQDPNNKSNVNFRECKQGPVWVSGVHHKLNTELWKFPESFKEGEQEVDLVNCRTN 896
QY 411 PQPTNLDSITAFWEGFHPETRPKLD-----EGSVLLH-----RTLGDKDAQRVQNLVSSLP 462
Db 897 EIIITGATVDFDGFEDFVNRLKNEKEKPMVLKLDKWPPEGEDPRDWMPSRFDLMANIP 956
QY 463 LPEYCAHQKLNLSYPLGLTLHLPLPQWAAVGVNSHRHGLCTKNLCVEVSDLSIL 521
Db 957 LPEYTRRDGKLNLSRFLNPFVVRPDLGPKMYNAYGLITTPEDRKYGTGTTNLHLDVSDAANVM 1016

QY 522 VHAEALPWPYRAQKDFLSGLDG-----EGLWSPGSGTSTVHVHVFRAQDAQRIR 570
Db 1017 VVVGIPKGCQEQEEVLRITIQDGSDELTIKRFIEGKEKPGA-----LWHIYAAKDKTEKIR 1072
QY 571 RFLQWVC-----PAGAGTLEPGAGSCYLDAGLRRLREMGVSCWTLLOAPGAVLVP 623
Db 1073 EFLKKVSEEQGQENPADH---DPIHDQSWYLDRLSKRLYQBYGQWAIQVFLGDDVFI 1129
QY 624 PAGAPHQVQGLVSTISVTHFLSPETSALSQAQCHQASLPPDH--RMLYAQMRAVFOA 681
Db 1130 PAGAPHQVHNLSCIKVAEDFVSPHVKHCFWLQFBRHLSQTHTNHEDKLQVKNVIYA 1189
QY 682 VKAAVQALQ 690
Db 1190 VKDAVAMLK 1198
RESULT 3
D86254
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D86254
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86254
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-851 <STO>
A:Cross-references: UNIPROT:O65384; UNIPARC:UPI00000A00D6; GB:AE005172; NID:g3157933; P
C:Genetics:
A:Map position: 1
Query Match 9.1%; Score 338; DB 2; Length 851;
Best Local Similarity 22.2%; Pred. No. 2.5e-16;
Matches 148; Conservative 91; Mismatches 192; Indels 236; Gaps 30;
QY 106 CSRCHHGLFNTHWRCSHCSHRLCV-ACGRIAG-----AGKREKTSQEQHTDD----- 153
Db 322 CNHCATSIVDLHRSCKPKSYELCNCCQEIREFPSSSVSEDTKTPSIKWNADENGSI 381
QY 154 CA-QEAGHAACSLILTOFVSSQALAEIETVMHQWAKFDIRGHCFCQVDARVWAPDGGQ 212
Db 382 CAPKELGGCDSVL-----ELKRILPVTWMS-DL-----E 410
QY 213 QKEPT-----EKTPTTPQPSGNGDSNRDKIETPDSTRESADGAGRSLPCPSLCEL 267
Db 411 QKAETFLASYSIKPPMSYCRCSDDMSMK-----RKAASRDGSSDNYLSPDSDLV 461
QY 268 LASTAVKLCGLGHDRIHMAFAPVTPALPSDDRIITNILDLSIIAQVVERKIQEALGPGLRAG 327
Db 462 LK-----QELL----- 468
QY 328 SGLRKGLSLPLSPVTRLSPPGALLWLQEPKPKHGFHLFOEHWROGQPVLV-SGIQKTLR 386
Db 469 -----H-FOEHWKSGEPVIVRNALNNTAG 491
QY 387 LS-----LWGMEL-----GTLCGQVQSLTALGPQPTNLDSITAFWEGFHPETR-----P 432
Db 492 LSWEPVWVW--RALCENVIDSAISSNMSDVKAIDCLANCEINTLCFFEGYSGKRTYENFWP 549
QY 433 KLDEGSVLLHLLHRTLGDKD--ASRVQNL-----VSSLPLPEYC-AHQGKLNLSYLP 480

550	EM	-----LKLQWPSPDKFENLLPHRCHDEFISALPQFVSDPRSGILNIATKLP	598
481	QY	LGLTLHPLEPOLWAAAYGVNSH--RGHLGTPKNLCVEVSDLSILVH-----ABA	526
599	DB	EGLLKPDLPKPTYYAIGTSDELGRGDSVTKLHC-DMSDAVNILMHTAEVTLSEEQRSATA	657
527	QY	QLPWPYRAQKD----FLSGLDGEGLWSP-----GSQTSIVHVFRAQDAQRIRRELQ---	574
658	DB	DLKQKHQOQNEKELQEQNGLEEEFVSDIIVVYDTSGLMDWDFKREDVPKLEEYLRKHC	717
575	QY	-----MVCPAGAGTLEPGAPGSCYLDAGLRRRLREEWG-----VSCWTLTLQAFGEAVLP	625
718	DB	IEFRHTYCSRVTKVYHPITHDQSYFLTVEHKRKLKAEFGMVTVIEWPTFVQKLGEAVFPA	777
626	QY	GAPHQVQGLVSTISVTOHFLSPETSALSQALCHOGASLPPDHRMLLYAOWDR-----	676
778	DB	GCPHQVRNLKCTKVAVDFVSPENTDECLRLTDFRQUPKNHK---AREDKLEIKQWVY	834
677	QY	AVFQAVK	683
835	DB	AVEQALK	841

RESULT 4

T01440
hypoetical protein F2401.2 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01440
R:Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Cor
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, January 1998
A:Description: Genomic sequence for Arabidopsis thaliana BAC F2401.
A:Reference number: Z14211
A:Accession: T01440
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-906 <SH1>
A:Cross-references: UNIPROT:O48794; UNIPARC:UPI000000A5BD0; EMBL:AC003113; NID:g2689438;
C:Genetics:
A:Gene: ATSP:F2401.2
A:Map position: 1
A:Introns: 239/2: 272/3: 317/1: 338/2: 682/3: 832/1: 860/3

Query Match 9.1%; Score 337; DB 2; Length 906;
Best Local Similarity 23.0%; Pred. No. 3.2e-16;
Matches 157; Conservative 79; Mismatches 240; Indels 2

Qy	89	EGQGFAMTDSFGPHC-----CSRCHGLFNTHWRCSCSHRLCVACGRIAGAGKNRK	143
Db	316	QGKLPEVEITAAISYTDERVYDCHCATSIVDLHRSPKCSYELCKCQOEIREGSLSER	375
Qy	144	TGSQEOHTDDCAQERAGHAACSILITQFVSSQALAE--LSTVMHQAWAKPDIRGHCFQCQVD	201
Db	376	PEMKPHYVD-----RGH-----RYMHGLDAAEPSLSTFDEBEAN-----PSD	413
Qy	202	ARVWAPGGQKQEPTEKTPPTPPQSCNGDSNRTKDIKEETPDSTESPAEDCAGRSPLPC	261
Db	414	AK-WSLGENSGITCAPEK-----LGGCGERMLELRILPLTWMSDLHKA-----	457
Qy	262	PSLCELLASTAVKLCGLGHDRIHMAFVPTPALPSDDRITNILDSTIAQVVERKIOEKALG	321
Db	458	-----ETFLSSYNISP-----RMLNCRCSSLELTETLTK-----	485
Qy	322	PGLRAGSLRGKLSLPSFVTRLS-----PGALLWLOEPPRPKGIFHLFQBHWKQGPV	376
Db	486	-----SASRTTSSDNYLFCPESLGLVKEEELH-----FOEHWAKGEVP	524
Qy	377	LV-SGIQKTLRLSLMGMEALGTGGQVOSLATGPPQPTNLDSTA-----FWEGF	425
Db	525	IVRNALDNTPGLSWEPVMWRALCENVNSTSSSEMSQVAKIDCLANCEVEINTROFFEGY	584
Qy	426	SHPETR-----PKLDEGSVILLHRTLTGDKDASRVQNIUVSSILPLPEYC-AHOG	471

[illegible]

RESULT 5

B85013
 hypothetical protein AT4g00990 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: B85013
 R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory
 Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: B85001; MUID:20083488; PMID:10617198
 A:Accession: B85013
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-730 <SPT>
 A:Cross-references: UNIPROT:Q9SV29; UNIPARC:UPI00000ABSE5; GB:NC_001268; NID:g7267596;
 C:Genetics:
 A:Gene: AT4g00990
 A:Map position: 4

Query Match 6.6%; Score 246; DB 2; Length 730;
Best Local Similarity 19.3%; Pred. No. 9e-10;

```

Best local similarity      100.0%, sized: no: 26-10,
Matches 154; Conservative 91; Mismatches 236; Indels 318; Gaps
Qy          91 QGPAMTSDSPGPHC-----CSRCHHGLFNTHWRC--SHCSHRLLCVACGRTAGAGK 139
Db          43 RGHPVTE--ANIKCKLDPSERIYCDLCRTSIANPHRSCPNKNCSVDICLSCKELSEG 100
Qy         140 NREKTGSQ--EQHTDDCAQEAGHAACSILITQFVSSQAALAEISTVMHQAWAKFDIRGHCF 177
Db         101 HQERDGKKAEGKGVECKRI PAGQGK----- 125
Qy         198 CQVDARVWAPCGGQQOKEPTKTPPTQPQSCNGDSNRKDIKEETPDSTESPAEDCAGRS 257
Db         126 ---DSDAYVPLHFSTWKLNSDSSIPCPPKKCGGGCTSTLEARR----- 165
Qy         258 PLPCPSICELASTAVKCLGH----DRIHMAFAPVTPALPDDRITNILDSIIAOVVE 312
Db         166 -LWKRDWVEKLITNAEKCTLNFRPTDVDIVHECSCSTNS-----DSIRRQAAF 213
Qy         313 RKIQEKALPGCLRAGSLRGKLSPLSPVRTRLSPGALLWLQEPKPHGFHLFOEHWMQ 372
Db         214 RKNANDNF-----LYSPNA-VDLAEDDIAH-----FGFHWMK 244
Qy         373 GQPVLVSGI-OKTLRLS-----LM-----GMEALGTGGOVSALTALGPPQPTN 415
Db         245 AEPVIVRNVLKTSGLSWEPVMVMRRACREMDPRKGTTEEITT---KVXALDCLDWCE-VE 300
Qy         416 LDSTAFAWEGF-----SHPETRPKLDEGSVLLLRHTLGDKDAASRVQNVLSSLPLPEYC 467

```

Db 301 INLHOFEGYLGEGRMHKGWPEMLKLDKOWPPSDLFKRLPRHNA-----EFIAALPFFDYT 356
QY 468 -AHQGLNLASYLPLGLTLHLPLEPOLWAAAYGVNS--HRGHLGTQNLCEVSDLSILVH- 523
Db 357 DPKSGILNLATRPFGESLKPDLGPKTYIAYGPHBELNRGDSVTKLHC-DISDAVNVLTHT 415
QY 524 AEAQLPP-----WYRAQKQFLSLDGEGL----- 547
Db 416 AKVEIPPVKYQNIKVHKKYAEAMLQKQYSGQVKEASELENKSMKEVDESKKLDKAA 475
QY 548 ----- 547
Db 476 NEEQNSNSRPSGSGAEKVIISKIARIELSHSVYVYKMLLNWENGLMPTLLATPPC 535
QY 548 -----WSPGSOTS-----TVMHVFRADQADQIRRF 572
Db 536 DTEEDNTPQAVSTSVESIQEQLDAPKETDGTGNTNERSKAVHGAVWDIFREDVPKLIQF 595
QY 573 LQ-----MVCAPAGATLBPAGPGCYLDAGLRRRLREBVGWVSCWTLQ 616
Db 596 LKRHEHEFRHFNNEPLESVIHPHDQTM-----FLSDSQKKQLKEBFDIEPWTFEQH 647
QY 617 PGEAVLPAGAPHOVQGLVSTISVTOHFLSPETSALSALCHOGASLPPDHRMLYAQMDR 676
Db 648 LGEAVFIPAGCPHQVNR---QVALDFVAPESVEECLRLTQEFRLPKDHS---SSEDK 700
QY 677 AVFQ--AVKAAVGALQEA 693
Db 701 LELKKIALYAASAIREV 719
RESULT 6
T10539
hypoetical protein F313.10 - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10539
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z17080
A:Accession: T10539
A:Molecule type: DNA
A:Residues: 1-730 <BEV>
A:Cross-references: UNIPROT:Q9SV29; UNIPARC:UPI00000AB5E5; EMBL:AL080237; GSPDB:GN00062;
A:Experimental source: cultivar Columbia; BAC clone F313
C:Genetics:
A:Gene: ATSP:F313.10
A:Map position: 4
A>Note: intron positions not resolved
Query Match 6.6%; Score 246; DB 2; Length 730;
Best Local Similarity 19.3%; Pred. No. 9e-10;
Matches 154; Conservative 91; Mismatches 236; Indels 318; Gaps 30;
QY 91 QGPAMTEDSPGIPHC-----CSCHHGLFNTHWRC--SHCSHRLCVACGRIAGAK 139
Db 43 RGHVTE--ANIKRKLDPSEIYCDLCRTSIANFHRSCPNKNCVSDICLSCKELSEG 100
QY 140 NREKTSQ--EQHTDDCAQAGHAACSLILTFQVSSQALAEISTVMHQAQWAKFDIRHCF 197
Db 101 HQBRDGKNAEGKGYECRIAGQK----- 125
QY 198 CQVDARVWAPGDGQOQKEPTKPTPTPQPSGNGDSNRTKDIKEETPDSTESPAEDGAGRS 257
Db 126 ---DSDAYVPLHFTWKLNSDSSIPCPKCEGGCGTSTLELR----- 165
QY 258 PLCPSPSCELLASTAVKLCIGH-----DRIHMAFAPVTPALPSPDDRITNILDIIAQVVE 312
Db 166 -LWKRDVVEKLITNAEKCCTLNFRPTDVIDIHECSCSCTNS-----DSIRQA 213
QY 313 RKIQEALGFCPLRAGSLRGLSLPLSPVRLTSPGALLWLQPRPKHGPHLFOEHW 372
Db 214 RKNADNF-----LYSPNA-VDLAEDDIAH-----FOFHWK 244

QY 373 GQPVLVSGI-OKTLRLS-----LW-----GMEALGTGCGVQSLTALGPPQPTN 415
Db 245 ABPTVTRNVLEKTSGLSWPEVMWRACREMDPKRGTEET---KVALDCLDWCE-VE 300
QY 416 LDSTAFWEGF-----SHPETRPKLDEGVSLLHLRTLGDKDSRVQNLVSLPLPEYC 467
Db 301 INLHOFEGYLGEGRMHKGWPEMLKLDKOWPPSDLFKRLPRHNA-----EFIAALPFFDYT 356
QY 468 -AHQGLNLASYLPLGLTLHLPLEPOLWAAAYGVNS--HRGHLGTQNLCEVSDLSILVH- 523
Db 357 DPKSGILNLATRPFGESLKPDLGPKTYIAYGPHBELNRGDSVTKLHC-DISDAVNVLTHT 415
QY 524 AEAQLPP-----WYRAQKQFLSLDGEGL----- 547
Db 416 AKVEIPPVKYQNIKVHKKYAEAMLQKQYSGQVKEASELENKSMKEVDESKKLDKAA 475
QY 548 ----- 547
Db 476 NEEQNSNSRPSGSGAEKVIISKIARIELSHSVYVYKMLLNWENGLMPTLLATPPC 535
QY 548 -----WSPGSOTS-----TVMHVFRADQADQIRRF 572
Db 536 DTEEDNTPQAVSTSVESIQEQLDAPKETDGTGNTNERSKAVHGAVWDIFREDVPKLIQF 595
QY 573 LQ-----MVCAPAGATLBPAGPGCYLDAGLRRRLREBVGWVSCWTLQ 616
Db 596 LKRHEHEFRHFNNEPLESVIHPHDQTM-----FLSDSQKKQLKEBFDIEPWTFEQH 647
QY 617 PGEAVLPAGAPHOVQGLVSTISVTOHFLSPETSALSALCHOGASLPPDHRMLYAQMDR 676
Db 648 LGEAVFIPAGCPHQVNR---QVALDFVAPESVEECLRLTQEFRLPKDHS---SSEDK 700
QY 677 AVFQ--AVKAAVGALQEA 693
Db 701 LELKKIALYAASAIREV 719
RESULT 7
T06461
DNA-binding protein P03, chloroplast - garden pea
C:Species: Pisum sativum (garden pea)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06461
R:Sato, N.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z15695
A:Accession: T06461
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1629 <SAT>
A:Cross-references: UNIPROT:O04698; UNIPARC:UPI000009FC7F; EMBL:X98744; NID:e995233; PI:1
A:Experimental source: cv. Alaska
C:Genetics:
A:Genome: nuclear
C:Keywords: chloroplast
Query Match 6.5%; Score 240.5; DB 2; Length 1629;
Best Local Similarity 19.1%; Pred. No. 6.2e-09;
Matches 157; Conservative 63; Mismatches 221; Indels 383; Gaps 27;
QY 106 CSCHHGLFNTHWRC--HCSHRLCVACG---RIAGAKNREKTSQEQHTDDCAQAGH 160
Db 934 CONCNTSIVNFHRSRCPNQCQYDLCTCTCTELRIGVHKDIPASGNEE----- 981
QY 161 AACSLILTFVSSQALAEISTVMHQAQWAKFDIRHCFQVDAVWAPGDGQOQKEPTK 220
Db 982 -----WVDA-----PPEGI 990
QY 221 PPTPQPSGNGDSNRTKDIKEETPDSTESP--ABDGAGRSPLPCPSICELLASTAVKLCIG 278
Db 991 P-----WRAETNGSI PCPPKARGGCGIATLSLRLE--ANWIDKLTGR 1032

Query Match 6.3%; Score 232; DB 2; Length 1701;
Best Local Similarity 19.6%; Pred. No. 2.7e-08;
Matches 171; Conservative 71; Mismatches 259; Indels 370; Gaps 33;
QY 71 RLCRLRKEREALAWAQREQ-----GPAMTEDSPGIPHC-----CSRCHHG 112
DB 947 KLFYLLKTKLPLLOHTQREQKSELEVASIHGSLMVEEKDILQAADDVDRVYCDNCNTS 1006
QY 113 LFETHWRC--SHGSHRLCVACGRIAGAKNREKTGSEQHTDDCAQAGHAACSLILITQF 170
DB 1007 IVNFHRSVNPYCRYDLCLTC-----CT----- 1029
QY 171 VSSQALAEISTVMQAWAKFDIGHCFQVDARWAPGDGGQKQKPTKPTPTPQPSGNG 230
DB 1030 -----ELRNGVHSHK-----DI-----PASGG--NEEMVNTPP----- 1054
QY 231 DSNRTKDIKEETPDSTESP--AEDGAGRSPLPCPSLCELL--ASTAVKLCIGHDRIHMAFA 287
DB 1055 ---ETIAWRAETNGSIPCPKARGCGGTATL---SURLFKANWIEKLTDRABELTIKYQ 1108
QY 288 PVTPALPDSDRITNILDSTIAQVVERKIOEALGPGRLAGSLRKLGLSLPLSPVTRLSP 347
DB 1109 P-----PIVDLSLECESECSFEEDA-----AHNSARKAASRETGHNDLLYCP 1150
QY 348 PGALLMLQEPKPKHGFHLFOEHWKQGPVLVSGIQK-----TLRLSLWME 393
DB 1151 DAIEIGDTE-----FDHFQHWIRGEPVIVRNYYKKGSLSDPMMWRAPFL----- 1198
QY 394 ALCTLGQVOSLTALGPPQ--PTNLDSTAFWEGF-----SHPETRPKLDGEGSVLLH 443
DB 1199 AKNILDADFTKADICLDWCEVQVNAFPFKGTYLTGRRYRNGWPEMLKLDMPNPNPFE 1258
QY 444 RTLGCDASRVQNLVSLPLPEYC-AHQGKLNILASVPLGLTLHP--LEPQLWAAVGV--N 499
DB 1259 DCLPRHCA---BFTAMLPDSYTHPKSGILNLATKLP--TVLKPDLPKTYIAYGALEB 1312
QY 500 SHRGHGTGNLCEVSDLSILVH-AEAQLPPWY----- 532
DB 1313 LSRGDSVTKLHC-DISDAVNILHTADVKTAPWQSKIHKKKYEVEDMRELYGLDSKA 1371
QY 533 ---RAOK----- 536
DB 1372 AGSRGRKRKRVRGVTVVDLKISEKEDINGRDSLTLLSEQEKDKLDREACVQEFSESTYSK 1431
QY 537 ----- 536
DB 1432 LDLNVSNOQVIDSPRFOQFDNLSDNLFVPRNDCESMLYDNVQRCSPRDCSGKNTS 1491
QY 537 -----DFLUGLGBGLWSPGSQT-----ST 556
DB 1492 VIDNQPCGGTKETTFVNGLSDSDISSSDIETDKIESVENEMPSNNLCGNDVHLETQYGSA 1551
QY 557 VWHVFRQAQRI-----RFLQWVCPAGAGTLEPGAGSCYLDAGLRRRLREWGV 608
DB 1552 VWDFRRQDVPKLEYLUNKHREFRHITSIPVNFVHIHQHDFYLNKHKHKKQLKLEYGV 1611
QY 609 SCWTLQAPGEAVLPVAPAGHQVGLVSTTSVTQHFLSPETSALSQACHQASLPDPDR 668
DB 1612 EPWTFEQLGEAVFIPAGCHQVNRKPCIKVAMDVFSPEVNVNCEVLTETEFRLLPKYHR 1671
QY 669 -----MLYAQMDRAVFOAQAAGA 688
DB 1672 SKEDKLEIKMALYA-ADVAIAEATK-LVGA 1700

RESULT 10
T05151
hypothetical protein F18E5.50 - Arabidopsis thaliana
N/Alternate names: hypothetical protein T6K22.160
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T05151; T05185
R/Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; New

submitted to the Protein Sequence Database, August 1998
A/Reference number: Z15400
A/Accession: T05151
A/Molecule type: DNA
A/Residues: 1-728 <BEV>
A/Cross-references: UNIPROT:O65408; UNIPARC:UPI00000A1A08; EMBL:AL022603
A/Experimental source: cultivar Columbia; BAC clone F18E5
A/Accession: T05185
A/Molecule type: DNA
A/Residues: 1-728 <BEW>
A/Cross-references: UNIPARC:UPI00000A1A08; EMBL:AL031187
A/Experimental source: cultivar Columbia; BAC clone T6K22
C/Genetics:
A/Map position: 4
A/Insertions: 208/3; 253/1; 298/1; 361/3; 471/3; 603/3; 628/1; 656/3; 696/3
A/Note: F18E5.50; T6K22.160
Query Match 5.5%; Score 205.5; DB 2; Length 728;
Best Local Similarity 19.8%; Pred. No. 7.4e-07;
Matches 121; Conservative 88; Mismatches 190; Indels 211; Gaps 25;
QY 220 TPPTPQSCNGSDSNRTKDIKEETPDSTESPADGAGRSPL-----PCPSLCE 266
DB 165 SPSP-----TTSNVSSPCDVKVGEEPISMIKRRFRSKNIEPLP-IGK 206
QY 267 LLASTAVKLCIGHDRI-HMAFA-----PVTPALPDSDRITNILDSTIAQVVERKIOEAL 320
DB 207 MODSQSVRSDI--DRVLHLHYAVCMLLPVLKEINAEBKVE-----VENDAERKVDLQRM 258
QY 321 GFCRLAGSGLR-----KGLSLPLSPVR--TRLSPGALLWLQEPKPKHGFHLFOEH 369
DB 259 --CTRSSVLRNLNSDQDSLSRKVGSKVCSNGIKSPKVLDDFHNNLEH----FQTH 312
QY 370 WRQGPVLVSGIQTKLRLSLWME-----LG 396
DB 313 WSKGHFVIVRSVSKSSSLNWDVPVLFCHYLMNRNNTGTTDCMDWFEVIGVKQFFLG 372
QY 397 TLGGQVQSITALGPPPOPTNLDSTAFWEGFSHPETRPKLDGEGSVLLHRTLGDKDSRVQN 456
DB 373 SLRGAETNTC-----QERLKL-EG--WLSSSLFKEQFPNHYAE 408
QY 457 LVSSSLPLPEYC-AHQGKLNILASVPLGLTLHPLEPQLWAAVGVNSHRGHL-----GTKNL 510
DB 409 ILNILPISHYMDPKRGLNLIAANLP--DTVQP--PDFGPCNLISYRSGBEYAPQDSVKKL 464
QY 511 CVEVSDLSILVH-----AEALPPWYRAQK-----D 537
DB 465 GFETCDMDVILLVYETPTVSTNQICRIRKLMKNIGVRSKNPAKGRSFRDKGKRDRLD 524
QY 538 FLUGLGE-----GLWSPGSQ-----TSTVWHVFRQAQRI 570
DB 525 DYSSSDSESSQHCLGAKRCRGEPEGEERESCNYSCEESLSNTYGAQWDYFQKQDVSKLL 584
QY 571 RF-----LQWVCPAGAGTLEPGAGSCYLDAGLRRRLREWGVSCWTLQAPGEAVLV 623
DB 585 EYIKNSHLESMDSSKKVSHPLLLQSYLYLDEYHKAUKKEEDFVPSFQCVGEAVIL 644
QY 624 PAGAPHQVGLVSTTSVTQHFLSPETSALSQACHQASLPDPDRMLYQMDRAVFOAVK 683
DB 645 PACFPQIRKNKSCVNAVLFKLSPEHVSSESIKRVKELNQLP-----QSVK 689
QY 684 AAVGALQEAQ 693
DB 690 SKANKIEVKK 699

RESULT 11
F86222
hypothetical protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: F86222
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Search completed: November 25, 2005, 22:14:21
Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 21:50:27 ; Search time 189 Seconds
(without alignments)
1611.056 Million cell updates/sec

Title: US-10-712-629b-18

Perfect score: 3704

Sequence: 1 VTQCSCVQAAGEVGLTGH.....MDRAVFQAVKAAGALQBAK 693

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3682	99.4	693	8	AD057501 Hairless
2	3682	99.4	1182	5	AAE19798 Mouse Hai
3	3682	99.4	1182	6	ABU62544 Mouse Hai
4	3510	94.8	1207	5	AAE19797 Rat Hairl
5	3510	94.8	1207	6	ABU62543 Rat Hairl
6	3095.5	83.6	1189	5	AAE19796 Human Hai
7	3095.5	83.6	1189	6	ABU62542 Human Hai
8	3093.5	83.5	1189	2	AAV15218 Human Hai
9	3089.5	83.4	1189	2	AAV15217 Human Hai
10	3089.5	83.4	1189	6	ABG72775 Human Hai
11	3088.5	83.4	984	5	AAE19794 Human Hai
12	3088.5	83.4	984	6	ABU62545 Human Hai
13	3077.5	83.1	1189	5	AAE19795 Human put
14	3077.5	83.1	1189	6	ABU62541 Human Hai
15	3067.5	82.8	984	6	ABU62540 Human Hai
16	7259	74.5	1134	8	ADQ80370 Protein h
17	7259	19.7	1265	7	ADJ68989 Human hea
18	728.5	19.7	1266	6	AAE19794 Human hea
19	728.5	19.7	1473	9	AAE20156 Novel hum
20	728.5	19.7	2055	8	ADB40509 Human nuc
21	700.5	18.9	1145	3	AAE42533 Human ORP
22	700.5	18.9	1417	7	ADJ70186 Human hea
23	618.5	16.7	179	6	ABU70953 Human adi
24	606.5	16.4	1236	7	ADD18797 Human dis

25	606.5	16.4	1321	8	AD019904	Ado19904 Human PRO
26	606.5	16.4	1321	9	ADX06345	Adx06345 Cyclin-de
27	606.5	16.4	1337	4	AM78460	Am78460 Human pro
28	606.5	16.4	1338	4	AM79444	Am79444 Human pro
29	606	16.4	671	5	ABF41598	Abf41598 Human ova
30	604.5	16.3	1214	7	ADD46411	Add46411 Rat Prote
31	603.5	16.3	1212	5	ABF65070	Abf65070 Hypoxia-i
32	603.5	16.3	1213	7	ADD46413	Add46413 Human pro
33	577.5	15.6	511	4	AAE41958	Aae41958 Human pol
34	577.5	15.6	511	9	AEA21016	Aea21016 Novel hum
35	550.5	14.9	854	4	ABE63729	AbE63729 Drosophil
36	448.5	12.1	1192	8	ADRI4639	Adri4639 Human NF-
37	362	9.8	875	8	ADO62965	Ado62965 Transcrip
38	337	9.1	906	8	ADO61723	Ado61723 Transcrip
39	309.5	8.4	917	8	ADO62489	Ado62489 Transcrip
40	305.5	8.2	860	8	ADO62492	Ado62492 Transcrip
41	305	8.2	963	9	ADW17337	Adw17337 Eucalyptu
42	297	8.0	747	8	ADX67858	Adx67858 Plant ful
43	297	8.0	1027	8	ADO61957	Ado61957 Transcrip
44	279.5	7.5	363	8	ADX73095	Adx73095 Plant ful
45	277.5	7.5	953	9	ADW18304	Adw18304 Eucalyptu

ALIGNMENTS

RESULT 1

AD057501

ID AD057501 standard; protein; 693 AA.

XX AC AD057501;

XX DT 29-JUL-2004 (first entry)

XX DE Hairless protein interaction partner #35.

XX mouse; hairless protein; Hrt; ubiquitous receptor UR; MAP1A;
KW KIAA0930 protein; monocytes antigen CD14; sphingolipid activator protein;
KW beta-eynuclen; C11 protein; vesicle-associated membrane protein 2;
KW aldolase A; CGL-106 protein; hypothalamus protein HSMNP1; alpha enolase;
KW POM-2P3; quinone oxidoreductase; pumilio 1; VPS41; KIAA0614 protein;
KW splicing factor CCL4; ubiquitin; beta-mannosidase; hair growth.
XX Mus musculus.
XX US2004086945-A1.
XX 06-MAY-2004.
XX 02-JUN-2003; 2003US-00452859.
XX 03-JUN-2002; 2002US-0385414P.
XX (PROC) PROCTER & GAMBLE CO.
XX Sreekrishna K, Gerwe GS, Toerner DR;
XX WPI; 2004-430095/40.
XX N-PSDB; AD057500.
XX New composition comprising mouse truncated hairless protein-human
XX interacting partner protein or nucleic acid complexes, useful for
XX screening test compounds that inhibit or enhance hair growth.
XX Example 1; SEQ ID NO 93; 60pp; English.
XX The invention relates to a composition comprising a mouse truncated
XX hairless (Hrt) protein-human interacting partner protein or nucleic acid
XX complex. The human interacting partner protein comprises a molecule
XX selected from ubiquitous receptor UR, MAP1A, KIAA0930 protein, monocytes
XX antigen CD14, sphingolipid activator protein, beta-synuclein, C11
XX protein, vesicle-associated membrane protein 2, aldolase A, CGL-106
XX protein, hypothalamus protein HSMNP1, alpha enolase, POM-2P3, quinone

CC oxidoreductase, pumilio 1, VPS41, KIAA0614 protein, splicing factor
CC CCl.4, ubiquitin, beta-mannosidase. Also described are: methods of
CC assaying a test compound for agonist or antagonist activity for the above
CC composition; and methods of inhibiting or increasing hair growth on a
CC surface in a subject. The present sequence represents a hairless protein
CC interacting partner used in the composition.
XX
SQ Sequence 693 AA;

Query Match 99.4%; Score 3682; DB 8; Length 693;
Best Local Similarity 99.4%; Pred. No. 7e-297; 3; Indels 0; Gaps 0;
Matches 689; Conservative 1; Mismatches 1

QY 1 VTQCQSCVQAAGVGLTGHSQKRSRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 60
DB 1 VTQCQSCVQAAGVGLTGHSQKRSRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 60
QY 61 AXHLLSGLGDRCLRLRKEREALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWRC 120
DB 61 AXHLLSGLGDRCLRLRKEREALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWRC 120
QY 121 SHCSHRLCVACGRIAGAGKKNREKTGSOEHTDDCAQEAAGHAACSLILTOFVSSQALAE 180
DB 121 SHCSHRLCVACGRIAGAGKKNREKTGSOEHTDDCAQEAAGHAACSLILTOFVSSQALAE 180
QY 181 TWMHQAWAKFDIRGHCFQVDARVWAPGDGGQKQKTEPTPTPPQPSGNGDSNRTKDIKE 240
DB 181 TWMHQAWAKFDIRGHCFQVDARVWAPGDGGQKQKTEPTPTPPQPSGNGDSNRTKDIKE 240
QY 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
DB 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
QY 301 NILDSIIAQVVERKIOEKALGPGLRAGSLRKGSLPLSPVTRRLSPPGALLWLQEPK 360
DB 301 NILDSIIAQVVERKIOEKALGPGLRAGSLRKGSLPLSPVTRRLSPPGALLWLQEPK 360
QY 361 HGFHLFQEHWRQGPVLVSGIQKTLRLSLWGMALGTLGGQVQSLTALGPPQPTNLDSTA 420
DB 361 HGFHLFQEHWRQGPVLVSGIQKTLRLSLWGMALGTLGGQVQSLTALGPPQPTNLDSTA 420
QY 421 FWEFGFHPETRPKLDDEGSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 480
DB 421 FWEFGFHPETRPKLDDEGSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 480
QY 481 LGTLHLPLPQLWAAVGVNSHRHGLGTKNLCVEVSDLSILVHAEAQLPWYRAQKDFLS 540
DB 481 LGTLHLPLPQLWAAVGVNSHRHGLGTKNLCVEVSDLSILVHAEAQLPWYRAQKDFLS 540
QY 541 GLDGEGLWSPGTSQTVWVHVFRAQDAQRIRRFQMVCPAGAGTLEPGAPGSCYLDAGLRR 600
DB 541 GLDGEGLWSPGTSQTVWVHVFRAQDAQRIRRFQMVCPAGAGTLEPGAPGSCYLDAGLRR 600
QY 601 RLREEVGVSCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQALCHQ 660
DB 601 RLREEVGVSCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQALCHQ 660
QY 661 ASLPPDHRLMYAQMDRAVFOAVKAAVGAALQEA 693
DB 661 ASLPPDHRLMYAQMDRAVFOAVKAAVGAALQEA 693

RESULT 2
AAE19798
ID AAE19798 standard; protein; 1182 AA.
XX
AC AAE19798;
XX
DT 18-JUN-2002 (first entry)
XX
DE Mouse Hairless protein (Hr).
XX
KW Mouse; Hairless protein; Hr protein; Hr gene; dermatological condition;

KW hair loss; gene therapy.
XX Mus musculus.
XX US6348348-B1.
XX PN 19-FEB-2002.
XX PD 07-APR-1999; 99US-00287354.
XX PF 07-APR-1998; 98US-0080888P.
XX PR (CARN-) CARNEGIE INST WASHINGTON.
XX PA Thompson CC;
XX PI WPI; 2002-204622/26.
XX DR Novel expression construct, useful in the diagnosis and treatment of
XX PT dermatological conditions, such as hair loss, contains a Hairless gene
XX PT sequence.
XX PS Example; Fig 1; 48pp; English.
XX PS The invention relates to human Hairless (Hr) polypeptides and nucleic
XX CC acid molecules (HR) encoding such polypeptides. The invention also
XX CC relates to the Hairless expression constructs which may be used in
XX CC transfection assays. Sequences of the invention are used in diagnosis
XX CC and treatment of dermatological conditions such as hair loss. They are
XX CC also used in gene therapy. Polynucleotides of the invention can be used
XX CC as probes for the detection of hair loss. The present sequence is mouse
XX CC Hairless protein (Hr)
XX SQ Sequence 1182 AA;

Query Match 99.4%; Score 3682; DB 5; Length 1182;
Best Local Similarity 99.4%; Pred. No. 1.5e-296;
Matches 689; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VTQCQSCVQAAGVGLTGHSQKRSRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 60
DB 490 VTQCQSCVQAAGVGLTGHSQKRSRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 549
QY 61 AXHLLSGLGDRCLRLRKEREALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWRC 120
DB 550 AXHLLSGLGDRCLRLRKEREALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWRC 609
QY 121 SHCSHRLCVACGRIAGAGKKNREKTGSOEHTDDCAQEAAGHAACSLILTOFVSSQALAE 180
DB 610 SHCSHRLCVACGRIAGAGKKNREKTGSOEHTDDCAQEAAGHAACSLILTOFVSSQALAE 669
QY 181 TWMHQAWAKFDIRGHCFQVDARVWAPGDGGQKQKTEPTPTPPQPSGNGDSNRTKDIKE 240
DB 670 TWMHQAWAKFDIRGHCFQVDARVWAPGDGGQKQKTEPTPTPPQPSGNGDSNRTKDIKE 729
QY 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
DB 730 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 789
QY 301 NILDSIIAQVVERKIOEKALGPGLRAGSLRKGSLPLSPVTRRLSPPGALLWLQEPK 360
DB 790 NILDSIIAQVVERKIOEKALGPGLRAGSLRKGSLPLSPVTRRLSPPGALLWLQEPK 849
QY 361 HGFHLFQEHWRQGPVLVSGIQKTLRLSLWGMALGTLGGQVQSLTALGPPQPTNLDSTA 420
DB 850 HGFHLFQEHWRQGPVLVSGIQKTLRLSLWGMALGTLGGQVQSLTALGPPQPTNLDSTA 909
QY 421 FWEFGFHPETRPKLDDEGSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 480
DB 910 FWEFGFHPETRPKLDDEGSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 969
QY 481 LGTLHLPLPQLWAAVGVNSHRHGLGTKNLCVEVSDLSILVHAEAQLPWYRAQKDFLS 540

Db 970 LGLTLHPLEPQLWAAAYGVNSHRHGLTKNLVCVEVSDLSILVHAEAQLPPWYRAQKDFLS 1029

QY 541 GLDGEGLWSPGSGTSTVWHVFRAQDAQRIRRFLOWVCPAGAGTLEPGAGSCVLDAGLRR 600

Db 1030 GLDGEGLWSPGSGTSTVWHVFRAQDAQRIRRFLOWVCPAGAGTLEPGAGSCVLDAGLRR 1089

QY 601 RLREEMGVSCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQAQLCHOG 660

Db 1090 RLREEMGVSCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQAQLYHOG 1149

QY 661 ASLPPDHRLMYAQMDRAVFAQVAAVGAALQAEAK 693

Db 1150 ASLPPDHRLMYAQMDRAVFAQVAAVGAALQAEAK 1182

RESULT 3

ABU62544

ID ABU62544 standard; protein; 1182 AA.

AC ABU62544;

XX

DT 08-SEP-2003 (first entry)

XX

DE Mouse Hairless (Hr) polypeptide.

XX

KW Mouse; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;

KW family pedigree; radiation hybrid; somatic cell hybrid.

XX

OS Mus sp.

XX

PN US2003027300-A1.

XX

PD 06-FEB-2003.

XX

PF 21-DEC-2001; 2001US-00024368.

XX

PR 07-APR-1998; 98US-0080888P.

PR 07-APR-1999; 99US-00287354.

XX

PA (CARN-) CARNEGIE INST WASHINGTON.

XX

PI Thompson CC;

XX

DR WPI; 2003-492034/58.

XX

PT New Hairless polypeptide and polynucleotide, useful for identifying and

PT detecting this genetic marker in family pedigrees or human-rodent somatic

PT cell hybrids, or detecting interacting proteins that bind hairless gene

PT or protein.

XX

PS Example; Fig 1; 34pp; English.

XX

CC The invention relates to a human hairless (HR) polynucleotide encoding

CC the hairless (Hr) polypeptide. The invention also relates to an

CC expression system comprising an expression construct which produces a

CC polypeptide with hairless transcription factor activity, a reporter

CC construct comprising a transcription regulatory region responsive to

CC hairless transcription activity to regulate transcription of the reporter

CC gene which is mediated by the transcription regulatory region and methods

CC of screening for chemical agents which modulate hairless-mediated

CC transcription, binding between hairless and thyroid hormone receptor or

CC hairless activity. The polynucleotide is useful as a probe or primer to

CC quantitate cognate RNA and DNA within cells, which can be subsequently

CC used to correlate hair growth or loss with hairless expression or

CC hairless-regulated transcription. The amino acid sequence of hairless

CC antigen can be used for preparing specific binding molecules (e.g.

CC polyclonal or monoclonal antibodies) for monitoring protein expression,

CC for affinity purification and for functional studies. The human hairless,

CC polynucleotide, polypeptide or specific binding molecule may be used to

CC identify and detect this genetic marker in family pedigrees, radiation

CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to

CC identify, isolate and detect interacting proteins that bind the hairless

CC gene or protein. This sequence represents a mouse hairless polypeptide of

CC the invention

XX

SQ Sequence 1182 AA;

Query Match 99.4%; Score 3682; DB 6; Length 1182;

Best Local Similarity 99.4%; Pred. No. 1.5e-296;

Matches 689; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VTQCSCVQAAGEVGLTGHSGKSRRSPLKEEKLEEDSSATSEEGGGPGPPASLUNKGL 60

Db 490 VTQCSCVQAAGEVGLTGHSGKSRRSPLKEEKLEEDSSATSEEGGGPGPPASLUNKGL 549

QY 61 AKHLLSGLGDRCLRLRKEREALAWAQREGQGQGPAMTEDSPGPHCCSRCHGLFNTHWRC 120

Db 550 AKHLLSGLGDRCLRLRKEREALAWAQREGQGQGPAMTEDSPGPHCCSRCHGLFNTHWRC 609

QY 121 SHCSHRLCVACGRIAGAKNREKTSQEOHTDDCAQEAHAACSLILTFVSSQALAEUS 180

Db 610 SHCSHRLCVACGRIAGAKNREKTSQEOHTDDCAQEAHAACSLILTFVSSQALAEUS 669

QY 181 TVMHOAWAKFDIRGHCFQCOVDARVWAPGGGQKQKPTPTPQPSGNGDSNRTKDIKE 240

Db 670 TVMHOAWAKFDIRGHCFQCOVDARVWAPGGGQKQKPTPTPQPSGNGDSNRTKDIKE 729

QY 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300

Db 730 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 789

QY 301 NILDSIIAQVVERKIQEAKLPGLRAGSLRGLSLPLSPVTRTRLSPPGALLWLQEPKPK 360

Db 790 NILDSIIAQVVERKIQEAKLPGLRAGSLRGLSLPLSPVTRTRLSPPGALLWLQEPKPK 849

QY 361 HGFHLFQEHWRGQPVLVSGIQKTLRLSLWGMEALGTGLGQVQSLTALGPPOPTNLDSTA 420

Db 850 HGFHLFQEHWRGQPVLVSGIQKTLRLSLWGMEALGTGLGQVQSLTALGPPOPTNLDSTA 909

QY 421 FWEGFSPETRPKLDGSGVLLLHRTLGLDKDASRVQNLVSSLPLPEYCAHQGLNLASYLP 480

Db 910 FWEGFSPETRPKLDGSGVLLLHRTLGLDKDASRVQNLVSSLPLPEYCAHQGLNLASYLP 969

QY 481 LGLTLHPLEPQLWAAAYGVNSHRHGLTKNLVCVEVSDLSILVHAEAQLPPWYRAQKDFLS 540

Db 970 LGLTLHPLEPQLWAAAYGVNSHRHGLTKNLVCVEVSDLSILVHAEAQLPPWYRAQKDFLS 1029

QY 541 GLDGEGLWSPGSGTSTVWHVFRAQDAQRIRRFLOWVCPAGAGTLEPGAGSCVLDAGLRR 600

Db 1030 GLDGEGLWSPGSGTSTVWHVFRAQDAQRIRRFLOWVCPAGAGTLEPGAGSCVLDAGLRR 1089

QY 601 RLREEMGVSCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQAQLCHOG 660

Db 1090 RLREEMGVSCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQAQLYHOG 1149

QY 661 ASLPPDHRLMYAQMDRAVFAQVAAVGAALQAEAK 693

Db 1150 ASLPPDHRLMYAQMDRAVFAQVAAVGAALQAEAK 1182

RESULT 4

AAE19797

ID AAE19797 standard; protein; 1207 AA.

XX

AC AAE19797;

XX

DT 18-JUN-2002 (first entry)

XX

DE Rat Hairless protein (Hr).

XX

KW Rat; Hairless protein; Hr protein; HR gene; dermatological condition;

KW hair loss; gene therapy.

XX

OS Rattus norvegicus.

XX

PN US6348348-B1.

XX 19-FEB-2002.
PD 07-APR-1999; 99US-00287354.
XX 07-APR-1998; 98US-0080888P.
PR (CARN-) CARNEGIE INST WASHINGTON.
XX Thompson CC;
PI WPI; 2002-204622/26.
XX Novel expression construct, useful in the diagnosis and treatment of
PT dermatological conditions, such as hair loss, contains a Hairless gene
PT sequence.
XX Example; Fig 1; 48pp; English.
XX The invention relates to human Hairless (Hr) polypeptides and nucleic
CC acid molecules (HR) encoding such polypeptides. The invention also
CC relates to the Hairless expression constructs which may be used in
CC transcription assays. Sequences of the invention are used in diagnosis
CC and treatment of dermatological conditions such as hair loss. They are
CC also used in gene therapy. Polynucleotides of the invention can be used
CC as probes for the detection of hair loss. The present sequence is rat
CC Hairless protein (Hr)
XX
XX Sequence 1207 AA;
SQ

Query Match 94.8%; Score 3510; DB 5; Length 1207;
Best Local Similarity 95.1%; Pred. No. 3.2e-282;
Matches 659; Conservative 11; Mismatches 21; Indels 2; Gaps 2;

QY 1 VTCQSCVQAGGVLGTGHSQSRSPLEEKQLEEDSSATSEEGGGGPGPEASLNKGL 60
DB 517 VTQPCSCVQAGGVEILTSHSQSHKPLPLEKPL-EEDSCATSEEGGS-SPPEASLNKGL 574
QY 61 AKHLLSGDLRLCLRLKEREALAWAQREGQGPAMTDSFGIPHCCSRCHHGLFNTHWRC 120
DB 575 AKHLLSGDLRLCLRLKEREALAWAQREGQGPAMTDSFGIPHCCSRCHHGLFNTHWRC 634
QY 121 SHCSHRLCVACGRIAGAGKREKTSQEQHTDCAQAGHAACSLILTVFSSQALAEIS 180
DB 635 SHCSHRLCVACGRIAGAGKREKTSQEQHTDCAQAGHAACSLILTVFSSQALAEIS 694
QY 181 TVNHQAWAKFDIRGHCFQVDARWAPCGGQOQKEPTKPTPPQSCNGDSNRTKDKE 240
DB 695 TVNHQAWAKFDIRGHCFQVDARWAPCGGQOQKEPTKPTPPQSCNGDSNRTKDKE 754
QY 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIIT 300
DB 755 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIIT 814
QY 301 NILDSIIAQVVERKIQEKALPGIRAGSLRGKLSLPLSVPTLSPGALLWLQEPKPK 360
DB 815 NILDSIIAQVVERKIQEKALPGIRAGSLRGKLSLPLSVPTLSPGALLWLQEPKPK 874
QY 361 HGPHLFQEHWRQOPVLVSGIKTLRLSLMGMEALGTGGQVQSLTALGPPQPTNLDSTA 420
DB 875 HGPHLFQEHWRQOPVLVSGIKTLRLSLMGMEALGTGGQVQSLTALGPPQPTNLDSTA 934
QY 421 FWGFGSHPEPRPKLDEGSVLLHRTLGCDASRVQNLVSSLPLPEYCAHQGKLNLSYLP 480
DB 935 FWGFGSHPEPRPKLDEGSVLLHRTLGCDASRVQNLVSSLPLPEYCAHQGKLNLSYLP 994
QY 481 LGTLTLPLEPQWAAVGVNSHRHGLTKNLCEVSDLSILVHAEALPQWYRAQKDFLS 540
DB 995 LGTLTLPLEPQWAAVGVNSHRHGLTKNLCEVSDLSILVHAEALPQWYRAQKDFLS 1054
QY 541 GLDGEGLWSPGSGTSTVWHVFRQAQDAQRIIRFLQWVCPAGAGTLEPGAPGSCYLDAGLR 600
DB 1055 GLDGEGLWSPGSGTSTVWHVFRQAQDAQRIIRFLQWVCPAGAGTLEPGAPGSCYLDAGLR 1114

QY 601 RLREEWGVSCTLLQAPGSAVLVPAGAPHQVQGLVSTISVTOHFLSPETSALSQALCHQG 660
DB 1115 RLREEWGVSCTLLQAPGSAVLVPAGAPHQVQGLVSTISVTOHFLSPETSALSQALCHQG 1174
QY 661 ASLPPDHRMLYAQMDRAVFAQVKAAYGALQAEK 693
DB 1175 ASLPPDHRMLYAQMDRAVFAQVKAAYGALQAEK 1207
RESULT 5
ABU62543
ID ABU62543 standard; protein; 1207 AA.
XX
XX AC ABU62543;
XX
DT 08-SEP-2003 (first entry)
XX
DE Rat Hairless (Hr) polypeptide.
XX
XX Rat; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;
KW family pedigree; radiation hybrid; somatic cell hybrid.
XX
XX Rattus sp.
OS
PN US2003027300-A1.
XX
XX 06-FEB-2003.
XX
XX 21-DEC-2001; 2001US-00024368.
XX
XX 07-APR-1998; 98US-0080888P.
PR 07-APR-1999; 99US-00287354.
XX
XX (CARN-) CARNEGIE INST WASHINGTON.
PA
PI Thompson CC;
XX
XX WPI; 2003-492034/58.
XX
XX New Hairless polypeptide and polynucleotide, useful for identifying and
PT detecting this genetic marker in family pedigrees or human-rodent somatic
PT cell hybrids, or detecting interacting proteins that bind hairless gene
PT or protein.
XX
XX Example; Fig 1; 34pp; English.
XX
XX The invention relates to a human hairless (HR) polynucleotide encoding
CC the hairless (Hr) polypeptide. The invention also relates to an
CC expression system comprising an expression construct which produces a
CC polypeptide with hairless transcription factor activity, a reporter
CC construct comprising a transcription regulatory region responsive to
CC hairless transcription activity to regulate transcription of the reporter
CC gene which is mediated by the transcription regulatory region and methods
CC of screening for chemical agents which modulate hairless-mediated
CC transcription, binding between hairless and thyroid hormone receptor or
CC hairless activity. The polynucleotide is useful as a probe or primer to
CC quantitate cognate RNA and DNA within cells, which can be subsequently
CC used to correlate hair growth or loss with hairless expression or
CC hairless-regulated transcription. The amino acid sequence of hairless
CC antigen can be used for preparing specific binding molecules (e.g.
CC polyclonal or monoclonal antibodies) for monitoring protein expression,
CC for affinity purification and for functional studies. The human hairless
CC polynucleotide, polypeptide or specific binding molecule may be used to
CC identify and detect this genetic marker in family pedigrees, radiation
CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to
CC identify, isolate and detect interacting proteins that bind the hairless
CC gene or protein. This sequence represents a rat hairless polypeptide of
CC the invention
XX
XX Sequence 1207 AA;
SQ

Query Match 94.8%; Score 3510; DB 6; Length 1207;

Best Local Similarity 95.1%; Pred. No. 3.2e-282;
Matches 659; Conservative 11; Mismatches 21; Indels 2; Gaps 2;
QY 1 VTQCSCVQAAGVGVLTGHSQKSRSPLEEKOLEEEDSSATSEEGGGPGPEASLNKGL 60
DB 517 VTQPCSCVQAGVEILTSHSQSHKLPLEEKPL-EEDSCATSEEGGGS-SPEASLNKGL 574
QY 61 AKHLLSGLGDRCLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCCSRCHHGLFNTHWRC 120
DB 575 AKHLLSGLGDRCLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCCSRCHHGLFNTHWRC 634
QY 121 SHCSHRLCVACGRIAGAKNREKTGSOEQHTDDCAQEAAGAACSLILTPFVSSQALAEIS 180
DB 635 SHCSHRLCVACGRIAGAKNREKTGSRQRTDDCAQEAAGAACSLILTPFVSSQALAEIS 694
QY 181 TVMHQAWAKFDIRGHCFQVDARVAPGDGGQKQTEKPTPTPOPSCHGDSNRTKDIXE 240
DB 695 TVMHQAWAKFDIRGHCFQVDARVAPGDGGQKQTEKPTPTPOPSCHGDSNRTKDIXE 754
QY 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRI 300
DB 755 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRI 814
QY 301 NILDSTIIAQVVERKIOEKALGPGLRAGSGRLKGLSLPLSPVTRTLPSPGALLWLQBRPK 360
DB 815 NILDSTIIAQVVERKIOEKALGPGLRAGSGRLKGLSLPLSPVTRTLPSPGALLWLQBRPK 874
QY 361 HGPHLFQEHWRQGPVLVSGIQKTLRLSLGMEALGTLGGQVOSLTALGPPQPTNLDSTA 420
DB 875 HGPHLFQEHWRQGPVLVSGIQKTLRLSLGMEALGTLGGQVOTLTALGPPQPTNLDSTA 934
QY 421 FWEFGFHPETRPKLDGSGVLLHRLTLDKXASRVQNLVSSLPPEYCAHQGKLNLAAYLP 480
DB 935 FWEFGFHPETRPKLDGSGVLLHRLTLDKXASRVQNLVSSLPPEYCAHQGKLNLAAYLP 994
QY 481 LGTLHPLFPQWAAAYGVNSHRHGLTKNLCVEVSDLSILVHAEAQLPWPYRAQKDFLS 540
DB 995 LGTLHPLFPQWAAAYGVNSHRHGLTKNLCVEVSDLSILVHAEAQLPWPYRAQKDFLS 1054
QY 541 GLDGEGLWSPGTSQTVMHVFRAQDAQRIIRRFLOWMPCPAGAGTLEPGAGSCYLDAGLR 600
DB 1055 GLDGEGLWSPGTSQTVMHVFRAQDAQRIIRRFLOWMPCPAGAGTLEPGAGSCYLDAGLR 1114
QY 601 LRREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSQAOLCHOG 660
DB 1115 LRREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSQAOLCHOG 1174
QY 661 ASLPPDHRMLYQMDRAVFAQVKAAYGALQEAQ 693
DB 1175 ASLPPDHRMLYQMDRAVFAQVKAAYGALQEAQ 1207

RESULT 6
AAE19796
ID AAE19796 standard; protein; 1189 AA.
XX
AC AAE19796;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human Hairless protein (Hr) #2.
XX
KW Human; Hairless protein; Hr protein; HR gene; dermatological condition;
KW hair loss; gene therapy.
XX
OS Homo sapiens.
XX
PN US6348348-B1.
XX
PD 19-FEB-2002.
XX
XX 07-APR-1999; 99US-00287354.
XX

PR 07-APR-1998; 98US-0080888P.
XX (CARN-) CARNEGIE INST WASHINGTON.
XX Thompson CC;
XX WPI; 2002-204622/26.
XX Novel expression construct, useful in the diagnosis and treatment of
XX dermatological conditions, such as hair loss, contains a Hairless gene
XX sequence.
XX Example; Fig 1; 48pp; English.
XX The invention relates to human Hairless (Hr) polypeptides and nucleic
XX acid molecules (HR) encoding such polypeptides. The invention also
XX relates to the Hairless expression constructs which may be used in
XX transcription assays. Sequences of the invention are used in diagnosis
XX and treatment of dermatological conditions such as hair loss. They are
XX also used in gene therapy. Polynucleotides of the invention can be used
XX as probes for the detection of hair loss. The present sequence is human
XX Hairless protein
XX
SQ Sequence 1189 AA;
Query Match 83.6%; Score 3095.5; DB 5; Length 1189;
Best Local Similarity 83.2%; Pred. No. 9.5e-248;
Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGVGVLTGHSQKSRSPLEEKOLEEEDSSATSEEGGGPGPEASLNKGL 59
DB 494 LAQCSCQAAGEGGGHACHSQVRRSPGLGELQEEEDTATNSSSEEGPGSPDSRLSTG 553
QY 60 LAKHLLSGLGDRCLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCCSRCHHGLFNTHWR 119
DB 554 LAKHLLSGLGDRCLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCCSRCHHGLFNTHWR 613
QY 120 CSCHSHRLCVACGRIAGAKNREKTGSOEQHTDDCAQEAAGAACSLILTPFVSSQALAE 179
DB 614 CPKCSHRLCVACGRIAGAKNREKTGSOEQHTDDCAQEAAGAACSLILTPFVSSQALAE 673
QY 180 STVMHQAQAFDIRGHCFQVDARVAPGDGGQKQTEKPTPTPOPSCHGDSNRTKDIX 239
DB 674 STAMHQAQAFDIRGHCFQVDARVAPGDGGQKQTEKPTPTPOPSCHGDSNRTKDIX 733
QY 240 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRI 299
DB 734 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRI 793
QY 300 TNILDSIIAQVVERKIOEKALGPGLRAGSGRLKGLSLPLSPVTRTLPSPGALLWLQBRPK 358
DB 794 TNILDSIIAQVVERKIOEKALGPGLRAGSGRLKGLSLPLSPVTRTLPSPGALLWLQBRPK 853
QY 359 -PKHGFHLFQEHWRQGPVLVSGIQKTLRLSLGMEALGTLGGQVOSLTALGPPQPTNLD 417
DB 854 CPKRGFHLFQEHWRQGPVLVSGIQKTLRLSLGMEALGTLGGQVOSLTALGPPQPTNLD 913
QY 418 STAFWEGFHPETRPKLDGSGVLLHRLTLDKXASRVQNLVSSLPPEYCAHQGKLNLA 477
DB 914 STTFWEGFHPETRPKLDGSGVLLHRLTLDKXASRVQNLVSSLPPEYCAHQGKLNLA 973
QY 478 YLPLGLTLPPLPQWAAAYGVNSHRHGLTKNLCVEVSDLSILVHAEAQLPWPYRAQK 537
DB 974 YLPPGLALPPLPQWAAAYGVNSHRHGLTKNLCVEVSDLSILVHAEAQLPWPYRAQK 1033
QY 538 FLGSLDGEGLWSPGTSQTVMHVFRAQDAQRIIRRFLOWMPCPAGAGTLEPGAGSCYLDAG 597
DB 1034 FLGSLDGEGLWSPGTSQTVMHVFRAQDAQRIIRRFLOWMPCPAGAGTLEPGAGSCYLDAG 1093
QY 598 LRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSQAOLC 657
DB 1094 LRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSQAOLC 1153

QY 658 HQGASLPDHRMLYAQMDFVFAVKAACALQAEK 693
DB 1154 HQGPSLPDCHLLYAQMDFVFAVKAACALQAEK 1189

RESULT 7
ABU62542
ID ABU62542 standard; protein; 1189 AA.

AC ABU62542;
DT 08-SEP-2003 (first entry)
XX Human Hairless (Hr) polypeptide #4.
XX Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;
KW family pedigree; radiation hybrid; somatic cell hybrid.
XX Homo sapiens.

OS
XX US2003027300-A1.

PN 06-FEB-2003.

PD 21-DEC-2001; 2001US-00024368.

PF 07-APR-1998; 98US-0080888P.

PR 07-APR-1999; 99US-00287354.

XX (CARN-) CARNEGIE INST WASHINGTON.

XX Thompson CC;

PI WPI; 2003-492034/58.

DR New Hairless polypeptide and polynucleotide, useful for identifying and
PT detecting this genetic marker in family pedigrees or human-rodent somatic
PT cell hybrids, or detecting interacting proteins that bind hairless gene
PT or protein.

XX Example; Fig 1; 34pp; English.

XX The invention relates to a human hairless (HR) polynucleotide encoding
CC the hairless (Hr) polypeptide. The invention also relates to an
CC expression system comprising an expression construct which produces a
CC polypeptide with hairless transcription factor activity, a reporter
CC construct comprising a transcription regulatory region responsive to
CC hairless transcription activity to regulate transcription of the reporter
CC gene which is mediated by the transcription regulatory region and methods
CC of screening for chemical agents which modulate hairless-mediated
CC transcription, binding between hairless and thyroid hormone receptor or
CC hairless activity. The polynucleotide is useful as a probe or primer to
CC quantitate cognate RNA and DNA within cells, which can be subsequently
CC used to correlate hair growth or loss with hairless expression or
CC hairless-regulated transcription. The amino acid sequence of hairless
CC antigen can be used for preparing specific binding molecules (e.g.
CC polyclonal or monoclonal antibodies) for monitoring protein expression,
CC for affinity purification and for functional studies. The human hairless
CC polynucleotide, polypeptide or specific binding molecule may be used to
CC identify and detect this genetic marker in family pedigrees, radiation
CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to
CC identify, isolate and detect interacting proteins that bind the hairless
CC gene or protein. This sequence represents a human hairless polypeptide of
CC the invention

XX Sequence 1189 AA;

Query Match 83.68; Score 3095.5; DB 6; Length 1189;
Best Local Similarity 83.21; Pred. No. 9.5e-248;
Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;

QY 1 VTCQSCVQAAGEVGLTGHISQKRRSPL-EEKQLEEDSSATSEEGGGGPGPEASLNKG 59

DB 494 LAQCSCAQAAGEGGGHACHSQVRRSPGLGELQEQEDTATNSSSEEGPGSPDRLSTG 553
QY 60 LAKHLLSGLDRLCRLLRKRERALAWAQREGGQGPAMTEDSPGPHCCSCHHGLFNTHWR 119
DB 554 LAKHLLSGLDRLCRLLRKRERALAWAQREGGQGPAMTEDSPGPHCCSCHHGLFNTHWR 613
QY 120 CSHCSHRLCVACGRIAGCKNREKTSQSOHTDDCAQEAAGHAAACSLILTFQVSSQALAE 179
DB 614 CPRCSHRLCVACGRVAGTGRAREKAGFQESABECTQEAGHAAACSLILTFQVSSQALAE 673
QY 180 STVMHQAQAFDIRGHCFQVDARVWAPGDGGQKEPTKPTPTPSCNGDSNRKDKIK 239
DB 674 STAMHQQVWKFDIRGHCFQVDARVWAPGDGGQKEPTKPTPTPSCNGDTHRTSIK 733
QY 240 EETPDSTESPDGAGRSPLPCPSICELLASTAVKLCGLGHDRIHMAFAPVTPALPDDRI 299
DB 734 EETPDSAETPAEDRAGRGLPCPSICELLASTAVKLCGLGHDRIHMAFAPVTPALPDDRI 793
QY 300 TNILDSIIAQVVERKIQEKALGPGLRAGSLRGLSLPLSPVTRTSLSPGALLWLQEP 358
DB 794 TNILDSIIAQVVERKIQEKALGPGLRAGSLRGLSLPLSPVTRTSLSPGALLWLQEP 853
QY 359 -PKHGFHLFQEHWRQGPVLVSGIQKTLRLSLMGMEALGTGCGVQSLTALGPPPTNLD 417
DB 854 CPERGFHLFQEHWRQGPVLVSGIQKTLRLSLMGMEALGTGCGVQSLTALGPPPTNLD 913
QY 418 STAFWEGFHPETRPKLDGSGVLLHRTLDGKASRVQNLVSSLPLPEYCAHOGKLNLAS 477
DB 914 STTFWEGFSPWELRPKSDGSGVLLHRTLDGKASRVQNLVSSLPLPEYCAHOGKLNLAS 973
QY 478 YLPLGLTLHPLPOLWAAVGVNSHRHGLGTKNLCVSVSLISLVHAEALQPPWYRAQKD 537
DB 974 YLPPGLALRPLPELWAAVGVNSHRHGLGTKNLCVSVSLISLVHAEALQPPWYRAQKD 1033
QY 538 FLSGLDGEGLWSPGSGTSTVWHVFRADQRIIRRFQWVCPAGAGTLEPGAGSCYLDAG 597
DB 1034 FLSGLDGEGLWSPGSGTSTVWHVFRADQRIIRRFQWVCPAGAGTLEPGAGSCYLDAG 1093
QY 598 LRRRLREWGVCWTLLOAPGEAVLPAGAPHQVGLVSTISVTQHFLSPETSALSQALC 657
DB 1094 LRRRLREWGVCWTLLOAPGEAVLPAGAPHQVGLVSTISVTQHFLSPETSALSQALC 1153

RESULT 8

ID AAY15218 standard; protein; 1189 AA.

XX AAY15218;

DT 09-NOV-1999 (first entry)

XX Human Hairless mutant amino acid sequence, Thr(1022)Ala.

XX alopecia; congenital alopecia; congenital atrichia;
KW androgenetic alopecia; alopecia areata; alopecia universalis; mutant;
KW hair follicle.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1022

FT /label= T1022A

XX WO9938965-A1.

XX 05-AUG-1999.

XX 29-JAN-1999; 99WO-US002128.

QY 60 LAKHLLSGLDRLCLRLRERELAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWR 119
Dd 554 LAKHLLSGLDRLCLRLRERELAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWR 613
QY 120 CSCHSHRLCVACGRIAGAGNREKTSQEOHTDDCAQEAHAACSLILTFVSSQALAE 179
Dd 614 CPRCSHRLCVACGRVAGTGRAREKAGFQEQSAEECTQEAHAACSLMLTFVSSQALAE 673
QY 180 STVMHQAWKFDIRGHCFQOVDARVWAPGDGQOKEPTKPTTPQSCNGDSNRKTDIK 239
Dd 674 STAMHQVWVKFDIRGHCPQOQADARVWAPGDAGQOKESTQKTPTTPQSCNGDTHRTKSIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 299
Dd 734 EETPDSAETPAEDRAGRGPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 793
QY 300 TNLDSIIAQVVERKIQEKALGRLAGSLRGLSLPLSPVTRLSPPGALLWLQEP - 358
Dd 794 TNLDSIIAQVVERKIQEKALGRLAGSLRGLSLPLSPVTRLSPPGALLWLQEPQP 853
QY 359 -PKHGFHLFOEHWROGQVVLVSGIQTKTLRLSLWGMALGTLGGVOSLTALGPPQPTNLD 417
Dd 854 CPRGFLHFOEHWROGQVVLVSGIQTKTLRLSLWGMALGTLGGVOSLTALGPPQPTNLD 913
QY 418 STAFWEGFHPETPKLDEGSVLLHRLTGLDGDASRVQNLVSSLPLPEYCAHOGKLNLAS 477
Dd 914 STTFWEGFHPETPKLDEGSVLLHRLTGLDGDASRVQNLVSSLPLPEYCAHOGKLNLAS 973
QY 478 YLPLGLTLHLEPOLWAAVGNVSHRGLTKNLCVEVDSLISILVHAEAOPLPWYRAQKD 537
Dd 974 YLPLGLTLHLEPOLWAAVGNVSHRGLTKNLCVEVDSLISILVHAEAOPLPWYRAQKD 1033
QY 538 FLGSLDGEGLWSPGSGTSTVWHVFRQAQDQIRIRFLQWCPAGAGTLEPGAGSCVLDAG 597
Dd 1034 FLGSLDGEGLWSPGSGTSTVWHVFRQAQDQIRIRFLQWCPAGAGTLEPGAGSCVLDAG 1093
QY 598 LRRLREWGVCWTLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSALC 657
Dd 1094 LRRLREWGVCWTLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSALC 1153
QY 658 HQGASLPDPHRLMYAQMDRAVFOAKAVALQAEK 693
Dd 1154 HQGASLPDPHRLMYAQMDRAVFOAKAVALQAEK 1189

RESULT 10
ID ABG72775 standard; protein; 1189 AA.

AC ABG72775;

DT 25-FEB-2003 (first entry)

XX Human hairless protein.

Human; catalytic DNA; catalytic RNA; hairless protein; hair loss;
atrachia; hair growth; hirsutism; catalytic nucleic acid; ribozyme;
DNazyme; self-catalytic; hammerhead ribozyme; deoxy-ribozyme;
catalytic core; cleavage site; pharmaceutical; hair production;
hair follicle; anagen phase; catagen phase; hair removal product;
depilatory.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 446

FT /note= "Encoded by TGG"

FT Misc-difference 584

FT /note= "Encoded by GGC"

XX W0200283891-A2.

XX 24-OCT-2002.

XX 12-APR-2002; 2002WO-US011683.
XX 13-APR-2001; 2001US-0283618P.
XX (UYCO) UNIV COLUMBIA NEW YORK.
XX Christiano AM;
XX WPI; 2003-093020/08.
XX N-PSDB; ABX13987.
XX New catalytic nucleic acid molecule that specifically cleaves Hairless
PT protein mRNA, useful for inhibiting hair production by a hair-producing
PT cell, or for inhibiting transition of a hair follicle from anagen phase
PT to catagen phase.
XX Disclosure; Page 59-63; 65pp; English.
XX

CC The invention discloses a new catalytic DNA or RNA molecule that
CC specifically cleaves, or inhibits expression of, Hairless protein mRNA
CC which comprises a catalytic domain that cleaves mRNA at a defined
CC consensus sequence and binding domains contiguous with the 5' and 3' ends
CC of the catalytic domain. Lack of expression of the hairless gene due to
CC inherited mutations leads to the complete loss of hair, known as
CC atrichia. Abundant hair growth, hirsutism, can be improved by inhibiting
CC the genes promoting hair growth, and one way to get targeted, transient
CC gene suppression is through the use of catalytic nucleic acid technology,
CC including ribozymes and DNazymes. Ribozymes are RNA structures which have
CC a self-catalytic enzymatic function and sequence specific RNA binding
CC ability. Small DNA oligonucleotides that have a similar structure to the
CC hammerhead ribozyme, called deoxy-ribozymes or DNazymes, having a
CC catalytic core and two sequence specific arms. The deoxy-ribozymes have
CC more lenient consensus cleavage site requirements and are less likely to
CC degrade in vivo, that hammerhead ribozymes. The catalytic nucleic acids
CC are useful in pharmaceutical compositions for inhibiting hair production
CC by a hair-producing cell, for inhibiting hair growth and for inhibiting
CC the transition of a hair follicle from the anagen phase to the catagen
CC phase. A non-human transgenic mammal is useful as a model for testing
CC hair removal products which function by inhibiting hairless protein
CC expression. The sequence presented is the human hairless protein
XX Sequence 1189 AA;

Query Match 83.4%; Score 3089.5; DB 6; Length 1189;
Best Local Similarity 83.0%; Pred. No. 3e-247;
Matches 578; Conservative 40; Mismatches 75; Indels 3; Gaps 2;

QY 1 VTQCSCVQAAGEVGVLTGHSQKSRSPLEEKQLEEDSSATSEGGGPGPEASLNKG 59
Dd 494 LAQCSCQAAGGEGGHACHSQQVRRSPGLGELQOEBEDATNSSSEGGSGPDSRLSTG 553
QY 60 LAKHLLSGLDRLCLRLRERELAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWR 119
Dd 554 LAKHLLSGLDRLCLRLRERELAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWR 613
QY 120 CSCHSHRLCVACGRIAGAGNREKTSQEOHTDDCAQEAHAACSLILTFVSSQALAE 179
Dd 614 CPRCSHRLCVACGRVAGTGRAREKAGFQEQSAEECTQEAHAACSLMLTFVSSQALAE 673
QY 180 STVMHQAWKFDIRGHCFQOVDARVWAPGDGQOKEPTKPTTPQSCNGDSNRKTDIK 239
Dd 674 STAMHQVWVKFDIRGHCPQOQADARVWAPGDAGQOKESTQKTPTTPQSCNGDTHRTKSIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 299
Dd 734 EETPDSAETPAEDRAGRGPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 793
QY 300 TNLDSIIAQVVERKIQEKALGRLAGSLRGLSLPLSPVTRLSPPGALLWLQEP - 358
Dd 794 TNLDSIIAQVVERKIQEKALGRLAGSLRGLSLPLSPVTRLSPPGALLWLQEPQP 853
QY 359 -PKHGFHLFOEHWROGQVVLVSGIQTKTLRLSLWGMALGTLGGVOSLTALGPPQPTNLD 417

Db 854 CPERGFHLFOEHWROQPVLVSGIORTLQGNLWGTALGALGQVQALSPGPPQSSLG 913
QY 418 STAFWEGFHPETPKLDEGSVLLHRTLGDKDASRVQNLVSLPLPEYCAHOGKLNLAS 477
Db 914 STTFWEGFSPWELPKSDSGSVLLHRTLGDKDASRVQNLVSLPLPEYCAHOGKLNLAS 973
QY 478 YLPLGLTLHPLEPOLWAAVGVNSHRHGLGTKNLCVEVSDLSILVHAEAQLPWPYRAQKD 537
Db 974 YLPPGLALRPLEPOLWAAVGVNSHRHGLGTKNLCVEVSDLSILVHAEAQLPWPYRAQKD 1033
QY 538 FLISGLDGEGLWSPGQSTVWHVFRADQRIIRRFLOWMCPAGAGTLEPCAGPSCYLDAG 597
Db 1034 FLISGLDGEGLWSPGQSTVWHVFRADQRIIRRFLOWMCPAGAGTLEPCAGPSCYLDAG 1093
QY 598 LRRRLREMGVSCWTLQAPGEAVLPAGAPHOVQGLVSTVSTQHFSLPETSALSQAOLC 657
Db 1094 LRRRLREMGVSCWTLQAPGEAVLPAGAPHOVQGLVSTVSTQHFSLPETSALSQAOLC 1153
QY 658 HQGASLPDPHRLMYAQMDRAVFAVKAAGALQAEK 693
Db 1154 HQGASLPDPHRLMYAQMDRAVFAVKAAGALQAEK 1189

RESULT 11

AAE19794
ID AAE19794 standard; protein; 984 AA.
XX
AC AAE19794;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human Hairless protein (Hr) #1.
XX
KW Human; Hairless protein; Hr protein; HR gene; dermatological condition;
KW hair loss; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 284
FT /note= "Encoded by CTT of the sequence shown as SEQ ID
NO:1 (AAD31072) in Column 35-38 of the specification"
FT Misc-difference 978
FT /note= "Encoded by CGG of the sequence shown as SEQ ID
NO:1 (AAD31077) in Column 25-32 of the specification"
FT Misc-difference 982
FT /note= "Encoded by CAG of the sequence shown as SEQ ID
NO:1 (AAD31077) in Column 25-32 of the specification"
XX
PN US6348348-B1.
XX
PD 19-FEB-2002.
XX
XX 07-APR-1999; 99US-00287354.
XX
PR 07-APR-1998; 98US-0080888P.
XX
PA (CARN-) CARNEGIE INST WASHINGTON.
XX
PI Thompson CC;
XX
XX WPI: 2002-204622/26.
DR N-PSDB; AAD31072, AAD31077.
XX
XX Novel expression construct, useful in the diagnosis and treatment of
PT dermatological conditions, such as hair loss, contains a Hairless gene
PT sequence.
XX
XX Claim 1b; Fig 1; 48pp; English.
XX
XX The invention relates to human Hairless (Hr) polypeptides and nucleic
CC acid molecules (HR) encoding such polypeptides. The invention also

CC relates to the Hairless expression constructs which may be used in
transcription assays. Sequences of the invention are used in diagnosis
CC and treatment of dermatological conditions such as hair loss. They are
CC also used in gene therapy. Polynucleotides of the invention can be used
CC as probes for the detection of hair loss. The present sequence is human
CC Hairless protein (Hr)
XX
SQ Sequence 984 AA;
Query Match 83.4%; Score 3088.5; DB 5; Length 984;
Best Local Similarity 83.0%; Pred. No. 2.8e-247;
Matches 578; Conservative 40; Mismatches 75; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGEVGLTGHQSQRSPLE-EBKQLEEDSSATSEBGGGPGGPEASLNKG 59
Db 289 LAQCSCQAAGEGGHACHSQVRRSPLEGGELQEBEDTATNSSSEGGPGGPDRLSTG 348
QY 60 LAKHLLSGLDRLCLRLRLKREALAWAQREGGQPMANEDSPGPHCCSCHHGLFNTHWR 119
Db 349 LAKHLLSGLDRLCLRLRLKREALAWAQREGGQPMANEDSPGPHCCSCHHGLFNTHWR 408
QY 120 CSHCSHRLCVACRTAGAKNREKTSQEQHTDDCAQEAHGAACSLILTFVSSQALAE 179
Db 409 CPRCSHRLCVACRVAGTGRAREKAGFQESAEBCTQEAHGAACSLILTFVSSQALAE 468
QY 180 STVMHQAWAKFDIRGHCFQVDARVWAPGDGGOKEPTKPTPTPOPCNGDSNRTKDIK 239
Db 469 STAMHQVWVKFDIRGHCFQVDARVWAPGDGGOKEPTKPTPTPOPCNGDTHRTKSIK 528
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPDDRI 299
Db 529 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPDDRI 588
QY 300 TNLDSIIAQVVERKIQEALGFLRAGSLRKLGLSLPLSPVTRLSPPGALLWLQEP- 358
Db 589 TNLDSIIAQVVERKIQEALGFLRAGSLRKLGLSLPLSPVTRLSPPGALLWLQEP 648
QY 359 -PKHGFHLFOEHWROQPVLVSGIORTLQGNLWGTALGALGQVQALSPGPPQSSLG 417
Db 649 CPERGFHLFOEHWROQPVLVSGIORTLQGNLWGTALGALGQVQALSPGPPQSSLG 708
QY 418 STAFWEGFHPETPKLDEGSVLLHRTLGDKDASRVQNLVSLPLPEYCAHOGKLNLAS 477
Db 709 STTFWEGFSPWELPKSDSGSVLLHRTLGDKDASRVQNLVSLPLPEYCAHOGKLNLAS 768
QY 478 YLPLGLTLHPLEPOLWAAVGVNSHRHGLGTKNLCVEVSDLSILVHAEAQLPWPYRAQKD 537
Db 769 YLPPGLALRPLEPOLWAAVGVNSHRHGLGTKNLCVEVSDLSILVHAEAQLPWPYRAQKD 828
QY 538 FLISGLDGEGLWSPGQSTVWHVFRADQRIIRRFLOWMCPAGAGTLEPCAGPSCYLDAG 597
Db 829 FLISGLDGEGLWSPGQSTVWHVFRADQRIIRRFLOWMCPAGAGTLEPCAGPSCYLDAG 888
QY 598 LRRRLREMGVSCWTLQAPGEAVLPAGAPHOVQGLVSTVSTQHFSLPETSALSQAOLC 657
Db 889 LRRRLREMGVSCWTLQAPGEAVLPAGAPHOVQGLVSTVSTQHFSLPETSALSQAOLC 948
QY 658 HQGASLPDPHRLMYAQMDRAVFAVKAAGALQAEK 693
Db 949 HQGASLPDPHRLMYAQMDRAVFAVKAAGALQAEK 984
RESULT 12
ABU62545
ID ABU62545 standard; protein; 984 AA.
XX
AC ABU62545;
XX
DT 08-SEP-2003 (first entry)
XX
DE Human Hairless (Hr) polypeptide #2.
XX
KW Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;


```
CC putative single zinc finger transcription factor protein (Hairless)
XX
SQ Sequence 1189 AA;

Query Match      83.1%; Score 3077.5; DB 5; Length 1189;
Best Local Similarity 82.9%; Pred. No. 3e-246;
Matches 577; Conservative 37; Mismatches 79; Indels 3; Gaps 2;

QY 1 VTQCSCVQAAGVGVLTHGSHQSRSPLE-EEKQLEEDSSATSEEGGGGPGPEASLNKG 59
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 494 LAQCSCQAAGGEGGHACHSQQVRSPGLGELQOEEDTATNSSEEGGSGDRLSTG 553
QY 60 LAKHLLSGLGDLRLCLRLKEREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHW 119
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 554 LAKHLLSGLGDLRLCLRLKEREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHW 613
QY 120 CSHCSHRLCVACGRVAGTGRAREKAGFQESAECECTQEAAGAACSLMLTQFVSSQALAE 179
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 614 CSHCSHRLCVACGRVAGTGRAREKAGFQESAECECTQEAAGAACSLMLTQFVSSQALAE 673
QY 180 STVMHQAQAKFDIRGHCFQVDARVWAPGDGQOKEPTKTPTTPOPCNGDSNRRTKDIK 239
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 674 STAHQVWVFKDIRGHCFQVDARVWAPGDGQOKEPTKTPTTPOPCNGDTHRTKSIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 299
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 734 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 793
QY 300 TNILDSIIAQVVERKIQEAKGLPGLRAGSLRGLSLPLSVTRLSPPGALLWLQEP 358
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 794 TNILDSIIAQVVERKIQEAKGLPGLRAGSLRGLSLPLSVTRLSPPGALLWLQEP 853
QY 359 -PKHGPHLFQEHWRQGPVLVSGIQKTLRLSLMGMEALGTGGQVQSLTALGPPQPTNLD 417
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 854 CPRGPHLFQEHWRQGPVLVSGIQKTLRLSLMGMEALGTGGQVQSLTALGPPQPTNLD 913
QY 418 STAFWEGFHPETPKLDEGVLHLLHRTLGDKDASRVQNLVSLPLPEYCAHQKLNLA 477
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 914 STTFWEGFHPETPKLDEGVLHLLHRTLGDKDASRVQNLVSLPLPEYCAHQKLNLA 973
QY 478 YLPLGLTLHLEPOLWAAVGNHSHGLGTHKNCVVEVDLSILVHAELPWWYRAQKD 537
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 974 YLPLGLTLHLEPOLWAAVGNHSHGLGTHKNCVVEVDLSILVHAELPWWYRAQKD 1033
QY 538 FLGSLDGEGLWSPGSGTSTVWHVFRQAQDQRIIRRFQWVCPAGAGTLEPGAGSCVLDAG 597
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 1034 FLGSLDGEGLWSPGSGTSTVWHVFRQAQDQRIIRRFQWVCPAGAGTLEPGAGSCVLDAG 1093
QY 598 LRRRLREEMGVSCWTLTQAFGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQALC 657
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 1094 LRRRLREEMGVSCWTLTQAFGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQALC 1153
QY 658 HQGASLPDPHMLYIAQMDRAVFAVAAVCAVGALEAK 693
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 1154 HQGASLPDPHMLYIAQMDRAVFAVAAVCAVGALEAK 1189

RESULT 14
ABU62541
ID ABU62541 standard; protein; 1189 AA.
XX
AC ABU62541;
XX
DT 08-SEP-2003 (first entry)
XX
DE Human Hairless (Hr) polypeptide #3.
XX
KW Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;
XX family pedigree; radiation hybrid; somatic cell hybrid.
XX
OS Homo sapiens.
XX
PN US2003027300-A1.
```

```
XX
PD
XX
PF 06-FEB-2003.
XX
XX 21-DEC-2001; 2001US-00024368.
XX
PR 07-APR-1998; 98US-0080888P.
XX 07-APR-1999; 99US-00287354.
XX
PA (CARN-) CARNEGIE INST WASHINGTON.
PI Thompson CC;
XX
XX WPI; 2003-492034/58.
XX
XX New Hairless polypeptide and polynucleotide, useful for identifying and
PT detecting this genetic marker in family pedigrees or human-rodent somatic
PT cell hybrids, or detecting interacting proteins that bind hairless gene
PT or protein.
XX
PS Example; Fig 1; 34pp; English.
XX
XX The invention relates to a human hairless (Hr) polynucleotide encoding
CC the hairless (Hr) polypeptide. The invention also relates to an
CC expression system comprising an expression construct which produces a
CC polypeptide with hairless transcription factor activity, a reporter
CC construct comprising a transcription regulatory region responsive to
CC hairless transcription activity to regulate transcription of the reporter
CC gene which is mediated by the transcription regulatory region and methods
CC of screening for chemical agents which modulate hairless-mediated
CC transcription, binding between hairless and thyroid hormone receptor or
CC hairless activity. The polynucleotide is useful as a probe or primer to
CC quantitate cognate RNA and DNA within cells, which can be subsequently
CC used to correlate hair growth or loss with hairless expression or
CC hairless-regulated transcription. The amino acid sequence of hairless
CC antigen can be used for preparing specific binding molecules (e.g.
CC polyclonal or monoclonal antibodies) for monitoring protein expression,
CC for affinity purification and for functional studies. The human hairless
CC polynucleotide, polypeptide or specific binding molecule may be used to
CC identify and detect this genetic marker in family pedigrees, radiation
CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to
CC identify, isolate and detect interacting proteins that bind the hairless
CC gene or protein. This sequence represents a human hairless polypeptide of
CC the invention
XX
SQ Sequence 1189 AA;

Query Match      83.1%; Score 3077.5; DB 6; Length 1189;
Best Local Similarity 82.9%; Pred. No. 3e-246;
Matches 577; Conservative 37; Mismatches 79; Indels 3; Gaps 2;

QY 1 VTQCSCVQAAGVGVLTHGSHQSRSPLE-EEKQLEEDSSATSEEGGGGPGPEASLNKG 59
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 494 LAQCSCQAAGGEGGHACHSQQVRSPGLGELQOEEDTATNSSEEGGSGDRLSTG 553
QY 60 LAKHLLSGLGDLRLCLRLKEREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHW 119
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 554 LAKHLLSGLGDLRLCLRLKEREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHW 613
QY 120 CSHCSHRLCVACGRVAGTGRAREKAGFQESAECECTQEAAGAACSLMLTQFVSSQALAE 179
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 614 CSHCSHRLCVACGRVAGTGRAREKAGFQESAECECTQEAAGAACSLMLTQFVSSQALAE 673
QY 180 STVMHQAQAKFDIRGHCFQVDARVWAPGDGQOKEPTKTPTTPOPCNGDSNRRTKDIK 239
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 674 STAHQVWVFKDIRGHCFQVDARVWAPGDGQOKEPTKTPTTPOPCNGDTHRTKSIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 299
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 734 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 793
QY 300 TNILDSIIAQVVERKIQEAKGLPGLRAGSLRGLSLPLSVTRLSPPGALLWLQEP 358
DQ : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 794 TNILDSIIAQVVERKIQEAKGLPGLRAGSLRGLSLPLSVTRLSPPGALLWLQEP 853
```

QY 359 -PKHGFHLFQEHWRQGPVLVSGIQKTLRLSLWGMALGTLGGQVQSLTALGPPQPTNLD 417
Db 854 CRRGFHLFQEHWRQGPVLVSGIQRTLQGNLWGTALGALGGQVQALSLAPPPSSLG 913
QY 418 STAFWGFSPHETRPKLDGSGVLLHRTLDGDKASRQNLVSSLPPLPEYCAHQKLNLS 477
Db 914 STTFWGFSPHETRPKLDGSGVLLHRTLDGDKASRQNLVSSLPPLPEYCAHQKLNLS 973
QY 478 YLPLGLTLHLEPQLWAAVGNVSHRGLGTNKLCEVSDLSILVHAEALPWPYRAQKD 537
Db 974 YLPLGLTLHLEPQLWAAVGNVSHRGLGTNKLCEVSDLSILVHAEALPWPYRAQKD 1033
QY 538 FLGLDGLWSPGSGTSTVHVFRQAQDQIRIRFLQWVCPAGAGTLEPGAPGSCYLDAG 597
Db 1034 FLGLDGLWSPGSGTSTVHVFRQAQDQIRIRFLQWVCPAGAGTLEPGAPGSCYLDAG 1093
QY 598 LRRRLREWGVSWTLLQAPGEAVLPAGAPHQVQGLVSTVTOHFLSPETSALSALQC 657
Db 1094 LRRRLREWGVSWTLLQAPGEAVLPAGAPHQVQGLVSTVTOHFLSPETSALSALQC 1153
QY 658 HQGASLPDPHRLMYAQMDFVFAVKAAGVCAQLEAK 693
Db 1154 HQGASLPDPHRLMYAQMDFVFAVKAAGVCAQLEAK 1189
RESULT 15
ID ABU62540 standard; protein; 984 AA.
AC ABU62540;
XX
XX 08-SEP-2003 (first entry)
XX Human Hairless (Hr) polypeptide #1.
XX Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;
KW family pedigree; radiation hybrid; somatic cell hybrid.
XX Homo sapiens.
FH Key Location/Qualifiers
FT Misc-difference 50 /note= "Encoded by ATC"
FT Misc-difference 140 /note= "Encoded by CAG"
FT Misc-difference 384 /note= "Encoded by CTG"
FT Misc-difference 446 /note= "Encoded by GAG"
FT Misc-difference 453 /note= "Residue printed as Her"
FT Misc-difference 467 /note= "Encoded by GAG"
FT Misc-difference 529..530 /note= "Encoded by GAGGAG"
FT Misc-difference 603 /note= "Encoded by AAC"
FT Misc-difference 606 /note= "Encoded by GAG"
FT Misc-difference 645 /note= "Encoded by GAG"
FT Misc-difference 649 /note= "Residue printed as Gys"
FT Misc-difference 659 /note= "Encoded by GAG"
FT Misc-difference 731 /note= "Residue printed as Leo"
FT Misc-difference 747 /note= "Encoded by AAG"
XX
XX US2003027300-A1.
XX

PD 06-FEB-2003.
XX 21-DEC-2001; 2001US-00024368.
XX 07-APR-1998; 98US-0080888P.
PR 07-APR-1999; 99US-00287354.
XX (CARN-) CARNEGIE INST WASHINGTON.
PA Thompson CC;
PI WPI: 2003-492034/58.
XX N-PSDB; ACD26351.
XX New Hairless polypeptide and polynucleotide, useful for identifying and
PT detecting this genetic marker in family pedigrees or human-rodent somatic
PT cell hybrids, or detecting interacting proteins that bind hairless gene
PT or protein.
XX
PS Claim 1; Page 14-17; 34pp; English.
XX The invention relates to a human hairless (HR) polynucleotide encoding
CC the hairless (Hr) polypeptide. The invention also relates to an
CC expression system comprising an expression construct which produces a
CC polypeptide with hairless transcription factor activity, a reporter
CC construct comprising a transcription regulatory region responsive to
CC hairless transcription activity to regulate transcription of the reporter
CC gene which is mediated by the transcription regulatory region and methods
CC of screening for chemical agents which modulate hairless-mediated
CC transcription, binding between hairless and thyroid hormone receptor or
CC hairless activity. The polynucleotide is useful as a probe or primer to
CC quantitate cognate RNA and DNA within cells, which can be subsequently
CC used to correlate hair growth or loss with hairless expression or
CC hairless-regulated transcription. The amino acid sequence of hairless
CC antigen can be used for preparing specific binding molecules (e.g.
CC polyclonal or monoclonal antibodies) for monitoring protein expression,
CC for affinity purification and for functional studies. The human hairless
CC polynucleotide, polypeptide or specific binding molecule may be used to
CC identify and detect this genetic marker in family pedigrees, radiation
CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to
CC identify, isolate and detect interacting proteins that bind the hairless
CC gene or protein. This sequence represents a human hairless polypeptide of
CC the invention
XX
XX Sequence 984 AA;
Query Match 82.8%; Score 3067.5; DB 6; Length 984;
Best Local Similarity 82.0%; Pred. No. 1.5e-245;
Matches 571; Conservative 47; Mismatches 75; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGVGVLTGHSQKRRSPL-EEKQLEEDSSATSBEGGGGPGPEASLNKG 59
Db 289 LAQCSCAQAAGGEGGHACHSQQVRRSPLEGELQOEEDTATNSSSEGGGSPDSRLSTG 348
QY 60 LAKHLGLGLDRLCLRLRKEREALAWAQEGQGPAMTEDSPGPHCCSRCHGLFNTHWR 119
Db 349 LAKHLGLGLDRLCLRLREREALAWAQEGQGPAMTGDSPGIPCCSRCHGLFNTHWR 408
QY 120 CSHCSHRLCVACGRIAGAGKNREKTSQEQHTDDCAQEAAGAACSLTLTQVSSQALAE 179
Db 409 CPRCSHRLCVACGRCVAGTGRAREKAGFQEQSAEECTQAGHAACSLMLTQVSSQALAE 468
QY 180 STVMHQAWAKFDIRGHCFQVDARVWAPGGGQKQKPTTEKTPTPQSCNGDSNRTDIK 239
Db 469 STAMHQVWVKFDIRGHCPQADARVWAPGDAGQKQKSTQKTPTPQSCNGDTHRTKSIK 528
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCGLHDIRHMAFPAPVTPALPSDDRI 299
Db 529 QQTPDPSAETPAEDRAGRGPUPCPSLCELLASTAVKLCGLHDIRHMAFPAPVTPALPSDDRI 588
QY 300 TNLIDSLIAQVVERKIQEALGFLRAGSLRKLGLPLSPVTRILSPPOALLWLQBPFR- 358
Db 589 TNLIDSLIAQVVERKIQKALGFLRAGSLRKLGLPLSPVTRILSPPOALLWLQBPQP 648

This Page Blank (Page 2)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 22:00:26 ; Search time 48 Seconds

(without alignments)
1193.630 Million cell updates/sec

Title: US-10-712-629B-18

Perfect score: 3704

Sequence: 1 VTQCSCVQAAGEVGLTGH.....MDRAVFQAVKAAGALQEAQ 693

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5.COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/6.COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/H.COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/PCUS.COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/RE.COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3682	99.4	1182	2	US-09-287-354-6
2	3510	94.8	1207	2	US-09-287-354-5
3	3095.5	83.6	1189	2	US-09-287-354-4
4	3095.5	83.6	1189	2	US-09-949-016-6931
5	3088.5	83.4	984	2	US-09-287-354-2
6	3077.5	83.1	1189	2	US-09-287-354-3
7	2333.5	63.0	943	2	US-09-949-016-7891
8	182	4.9	177	2	US-09-640-211A-1138
9	152.5	4.1	2442	2	US-09-514-247A-10
10	152.5	4.1	2442	2	US-09-538-092-1370
11	146.5	4.0	2414	1	US-08-227-536-2
12	146.5	4.0	2414	1	US-09-538-092-1289
13	146.5	4.0	2414	4	PCT-US95-04682-2
14	145.5	3.9	2441	1	US-08-194-468-2
15	145.5	3.9	2441	2	US-08-961-739-2
16	145.5	3.9	2441	2	US-09-514-247A-8
17	145.5	3.9	2441	2	US-09-686-316-2
18	127.5	3.4	1479	2	US-08-840-062-2
19	124.5	3.4	427	2	US-09-502-540-10191
20	121	3.3	732	1	US-08-317-522A-5
21	121	3.3	778	1	US-08-439-818A-5
22	121	3.3	778	1	US-08-751-965-5
23	121	3.3	778	1	US-08-738-975-5
24	121	3.3	778	1	US-08-728-626-5
25	121	3.3	778	2	US-08-808-599A-5
26	121	3.3	987	2	US-09-252-991A-27483
27	120	3.2	861	2	US-09-252-991A-18375

28	118	3.2	656	2	US-09-902-540-12404	Sequence 12404, A
29	117.5	3.2	662	2	US-10-104-047-3398	Sequence 3398, Ap
30	116.5	3.1	2157	2	US-09-466-778-2	Sequence 2, Appli
31	115.5	3.1	574	2	US-09-079-431B-6	Sequence 6, Appli
32	115.5	3.1	609	2	US-09-252-991A-32771	Sequence 32771, A
33	115	3.1	715	2	US-09-252-991A-32740	Sequence 32740, A
34	113.5	3.1	739	2	US-09-252-991A-32778	Sequence 32778, A
35	113.5	3.1	2509	1	US-08-469-005A-10	Sequence 10, Appli
36	112.5	3.0	1208	2	US-09-463-702A-2	Sequence 2, Appli
37	112.5	3.0	1208	2	US-09-699-135-2	Sequence 2, Appli
38	112.5	3.0	2418	2	US-09-949-016-10703	Sequence 10703, A
39	112.5	3.0	2511	2	US-09-261-907-2	Sequence 2, Appli
40	111.5	3.0	442	1	US-08-208-108-2	Sequence 2, Appli
41	110.5	3.0	683	2	US-09-949-016-11117	Sequence 11117, A
42	110.5	3.0	1147	1	US-08-131-365B-38	Sequence 38, Appli
43	110.5	3.0	1147	1	US-08-668-123-38	Sequence 38, Appli
44	110.5	3.0	1164	2	US-09-949-016-9845	Sequence 9845, Ap
45	110	3.0	386	2	US-09-605-703B-2514	Sequence 2514, Ap

ALIGNMENTS

RESULT 1

US-09-287-354-6

; Sequence 6, Application US/09287354

; Patent No. 6348348

; GENERAL INFORMATION:

; APPLICANT: THOMPSON, Catherine C.

; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN

; FILE REFERENCE: Thompson-20263/0243435

; CURRENT APPLICATION NUMBER: US/09/287,354

; CURRENT FILING DATE: 1999-04-07

; EARLIER APPLICATION NUMBER: US 60/080,888

; EARLIER FILING DATE: 1998-04-07

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6

; LENGTH: 1182

; TYPE: PRT

; ORGANISM: Mouse

; US-09-287-354-6

Query Match 99.4%; Score 3682; DB 2; Length 1182;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 689; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY	1	VTQCSCVQAAGEVGLTGH	SQKRRSPLEEKLEEDSSATSEEGGGPGPEASLNKGL	60
DB	490	VTQCSCVQAAGEVGLTGH	SQKRRSPLEEKLEEDSSATSEEGGGPGPEASLNKGL	549
QY	61	AKHLISGLGDRICLLRKRE	ALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWRC	120
DB	550	AKHLISGLGDRICLLRKRE	ALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWRC	609
QY	121	SHCSHRLCVACGRIAGAG	KNRKTSQEOHTDDCAQEAHAACSLILTOPVSSQALAEIS	180
DB	610	SHCSHRLCVACGRIAGAG	KNRKTSQEOHTDDCAQEAHAACSLILTOPVSSQALAEIS	669
QY	181	TYMHQAWAKFDIRGHCF	QCVDAVWAPGGGQKQKTEPTPTPPQPSGNSNRDKIKE	240
DB	670	TYMHQAWAKFDIRGHCF	QCVDAVWAPGGGQKQKTEPTPTPPQPSGNSNRDKIKE	729
QY	241	ETPDSTESPAEDGAGRS	PLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT	300
DB	730	ETPDSTESPAEDGAGRS	PLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT	789
QY	301	NILDSTIAOVVERKIOE	KALGRLAGSLPLSPVTRLSPPGALLWLQSPRPK	360
DB	790	NILDSTIAOVVERKIOE	KALGRLAGSLPLSPVTRLSPPGALLWLQSPRPK	849
QY	361	HGFHLFOEHRWQOPVL	VSGIQKTLRLSLWGMEALGTLCGQVQSITALGPQPTNLDSTA	420

Db 850 HGFRLQEHWRQGPVLVSGIQKTLRLSLWMEALGTLGGQVQTLTALGPPQPTNLDSTA 909
QY 421 FMEGFSHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHQGKLNLSYLP 480
Db 910 FMEGFSHPETRPKLDGSGVLLHRTLGDKDASRVQNLASSLPPEYCAHQGKLNLSYLP 969
QY 481 LGLTLHPLPQLWAAAGVNSHRHGLGTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLS 540
Db 970 LGLTLHPLPQLWAAAGVNSHRHGLGTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLS 1029
QY 541 GLDGEGLWSPGSGTSTVWHVFRQAQDAQRIIRFLQMVCPAGAGTLEPGAPGSCYLDAGLRR 600
Db 1030 GLDGEGLWSPGSGTSTVWHVFRQAQDAQRIIRFLQMVCPAGAGTLEPGAPGSCYLDAGLRR 1089
QY 601 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQAQLCHQG 660
Db 1090 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQAQLCHQG 1149
QY 661 ASLPPDHRMLYAQMDDRAVFOAVKAAVGAALQEA 693
Db 1150 ASLPPDHRMLYAQMDDRAVFOAVKAAVGAALQEA 1182

RESULT 2

US-09-287-354-5
; Sequence 5, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: Rat
US-09-287-354-5

Query Match 94.8%; Score 3510; DB 2; Length 1207;

Best Local Similarity 95.1%; Pred. No. 0;

Matches 659; Conservative 11; Mismatches 21; Indels 2; Gaps 2;

QY 1 VTQCSCVQAAGEVGLTGHSSQKSRRLPBEKOLEEEDSSATSEGGGGPPEASLNKGL 60
Db 517 VTQCSCVQAAGEVEILTSHSQSKHLPLBEKPL-EEDSCATSEGGGS-SPEASLNKGL 574
QY 61 AKHLLSGLGDRCLRLRKEREALAWAQREGQGPAMTEDSPGIPHCSCRCHHGLFNTHWR 120
Db 575 AKHLLSGLGDRCLRLRKEREALAWAQREGQGPAMTEDSPGIPHCSCRCHHGLFNTHWR 634
QY 121 SHCSHRLCVACGRITAGAKNREKTSQEQHTDDCAQAGHAACSLIITQFVSSQALAE 180
Db 635 SHCSHRLCVACGRITAGAKNREKTSQEQHTDDCAQAGHAACSLIITQFVSSQALAE 694
QY 181 TVMHOAWAKFDIRGHCFQVDARVWAPGDGQKQKTEKPTPTPQPCNGSDNRTKDIKE 240
Db 695 TVMHOAWAKFDIRGHCFQVDARVWAPGDGQKQKTEKPTPTPQPCNGSDNRTKDIKE 754
QY 241 ETPDSTESPAEDGAGRSPLPCPSICELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
Db 755 ETPDSTESPAEDGAGRSPLPCPSICELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 814
QY 301 NILDSIIAQVVERKIQEALGPGLRAGSLRKLGLSLPLSPVTRLSPGALLWLQEPK 360
Db 815 NILDSIIAQVVERKIQEALGPGLRAGSLRKLGLSLPLSPVTRLSPGALLWLQEPK 874
QY 361 HGFHLQEHWRQGPVLVSGIQKTLRLSLWMEALGTLGGQVQSLTALGPPQPTNLDSTA 420

Db 875 HGFRLQEHWRQGPVLVSGIQKTLRLSLWMEALGTLGGQVQTLTALGPPQPTSLDSTA 934
QY 421 FMEGFSHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHQGKLNLSYLP 480
Db 935 FMEGFSHPETRPKLDGSGVLLHRTLGDKDASRVQNLASSLPPEYCAHQGKLNLSYLP 994
QY 481 LGLTLHPLPQLWAAAGVNSHRHGLGTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLS 540
Db 995 LGLTLHPLPQLWAAAGVNSHRHGLGTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLS 1054
QY 541 GLDGEGLWSPGSGTSTVWHVFRQAQDAQRIIRFLQMVCPAGAGTLEPGAPGSCYLDAGLRR 600
Db 1055 GLDGEGLWSPGSGTSTVWHVFRQAQDAQRIIRFLQMVCPAGAGTLEPGAPGSCYLDAGLRR 1114
QY 601 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQAQLCHQG 660
Db 1115 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQAQLCHQG 1174
QY 661 ASLPPDHRMLYAQMDDRAVFOAVKAAVGAALQEA 693
Db 1175 ASLPPDHRMLYAQMDDRAVFOAVKAAVGAALQEA 1207

RESULT 3

US-09-287-354-4
; Sequence 4, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-287-354-4

Query Match 83.6%; Score 3095.5; DB 2; Length 1189;

Best Local Similarity 83.2%; Pred. No. 1.7e-282;

Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;

QY 1 VTQCSCVQAAGEVGLTGHSSQKSRRLP-EKQLEEEEDSSATSEGGGGPPEASLNK 59
Db 494 LAQCSCAQAAGGGGACHSQQVRSPLGGELQEEEDTATNSSSEEGPGSPDRSLSTG 553
QY 60 LAKHLLSGLGDRCLRLRKEREALAWAQREGQGPAMTEDSPGIPHCSCRCHHGLFNTHWR 119
Db 554 LAKHLLSGLGDRCLRLRKEREALAWAQREGQGPAMTEDSPGIPHCSCRCHHGLFNTHWR 613
QY 120 CSFCSHRLCVACGRITAGAKNREKTSQEQHTDDCAQAGHAACSLIITQFVSSQALAE 179
Db 614 CPRCSHRLCVACGRVAGTGRAREKAGFQEQSECTQEAGHAACSLMLTQFVSSQALAE 673
QY 180 STVMHOAWAKFDIRGHCFQVDARVWAPGDGQKQKTEKPTPTPQPCNGSDNRTKDIK 239
Db 674 STAHQVWAKFDIRGHCFQVDARVWAPGDGQKQKTEKPTPTPQPCNGDTHRTKSIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSICELLASTAVKLCIGHDRIHMAFAPVTPALPSDDR 299
Db 734 EETPDSAETPAEDRAGRGLPCPSICELLASTAVKLCIGHDRIHMAFAPVTPALPSDDR 793
QY 300 TNLDSIIAQVVERKIQEALGPGLRAGSLRKLGLSLPLSPVTRLSPGALLWLQEPK 358
Db 794 TNLDSIIAQVVERKIQEALGPGLRAGSLRKLGLSLPLSPVTRLSPGALLWLQEPK 853
QY 359 -PKHGFHLQEHWRQGPVLVSGIQKTLRLSLWMEALGTLGGQVQSLTALGPPQPTNLD 417

Db 854 CPRRGPHLFOEHWROQOPVLVSGIQRITLQGNLWTEALGALGGQVQALSPLGPPQPSLSG 913
QY 418 STAFWEGFSPETPKLDEGSVLLHRTLGDKDASRYONLVSSILPLPEYCAHQKUNLAS 477
Db 914 STTFWEGFSPPELKPCKDEGSVLLHRTLGDKDASRYONLVSSILPLPEYCAHQKUNLAS 973
QY 478 YLPGLGLTLHPLEPOLMAAYGVNSHRHGLTKNLCVEVSDILSILVHAEALPMPYRAQKD 537
Db 974 YLPGLGLTLHPLEPOLMAAYGVNSHRHGLTKNLCVEVSDILSILVHAEALPMPYRAQKD 1033
QY 538 FLSLDDEGLWSPGSGTSTVHVHFRQAQDQRIIRFLQWCPAGAGTLEPGAGSGCYLDAG 597
Db 1034 FLSLDDEGLWSPGSGTSTVHVHFRQAQDQRIIRFLQWCPAGAGTLEPGAGSGCYLDAG 1093
QY 598 LRRRLREMGVSCWTLLOAPGEAVLVPAGAPHOVQGLVSTISVTQHFLSPETSALSALC 657
Db 1094 LRRRLREMGVSCWTLLOAPGEAVLVPAGAPHOVQGLVSTISVTQHFLSPETSALSALC 1153
QY 658 HQGASLPDPHRLMYAQMDRAVFOAVKAAVAGALQEA 693
Db 1154 HQGASLPDPHRLMYAQMDRAVFOAVKAAVAGALQEA 1189

RESULT 4

US-09-949-016-6931
; Sequence 6931, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6931
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6931

Query Match 83.6%; Score 3095.5; DB 2; Length 1189;
Best Local Similarity 83.2%; Pred. No. 1.7e-282;
Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGEVGLTGHKSRRSPLEKQLEEDSSATSEGGGPGPEASLNGK 59
Db 494 LAQCSCQAAGGGGCHACHSQVRRSPGLGELQEBEDTATNSSEEGPGSGDRLSTG 553
QY 60 LAKHLLSGLDRLCLLRLKEREALAWAQREGQGPANTEPGPHCCSRCHGLFNTHWR 119
Db 554 LAKHLLSGLDRLCLLRLKEREALAWAQREGQGPANTEPGPHCCSRCHGLFNTHWR 613
QY 120 CSHCSHRLCVACGRVAGTGRAREKAGFQESAECTQEAAGAACSLMLTQFVSSQALAE 179
Db 614 CPRCSHRLCVACGRVAGTGRAREKAGFQESAECTQEAAGAACSLMLTQFVSSQALAE 673
QY 180 STVHQAWAKFDIRGHCFQVDARVWAPGGGQKQEPTEKTPPTPQSCNGDSNRKDK 239
Db 674 STAHQVWVWFDIRGHCFQVDARVWAPGGGQKQEPTEKTPPTPQSCNGDTHRTKSIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRTHMAFAPVTPALPSDDRI 299
Db 734 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRTHMAFAPVTPALPSDDRI 793
QY 300 TNLDSIIAQVVERKIQEALGPGLRAGSLRKLGLSLPSVTRLSPPGALLWLQEP - 358

Db 794 TNLDSIIAQVVERKIQEALGPGLRAGSLRKLGLSLPSVTRLSPPGALLWLQEP - 853
QY 359 -PKHGFHLFOEHWROQOPVLVSGIQRITLQGNLWTEALGALGGQVQALSPLGPPQPSLSG 417
Db 854 CPRRGPHLFOEHWROQOPVLVSGIQRITLQGNLWTEALGALGGQVQALSPLGPPQPSLSG 913
QY 418 STAFWEGFSPETPKLDEGSVLLHRTLGDKDASRYONLVSSILPLPEYCAHQKUNLAS 477
Db 914 STTFWEGFSPPELKPCKDEGSVLLHRTLGDKDASRYONLVSSILPLPEYCAHQKUNLAS 973
QY 478 YLPGLGLTLHPLEPOLMAAYGVNSHRHGLTKNLCVEVSDILSILVHAEALPMPYRAQKD 537
Db 974 YLPGLGLTLHPLEPOLMAAYGVNSHRHGLTKNLCVEVSDILSILVHAEALPMPYRAQKD 1033
QY 538 FLSLDDEGLWSPGSGTSTVHVHFRQAQDQRIIRFLQWCPAGAGTLEPGAGSGCYLDAG 597
Db 1034 FLSLDDEGLWSPGSGTSTVHVHFRQAQDQRIIRFLQWCPAGAGTLEPGAGSGCYLDAG 1093
QY 598 LRRRLREMGVSCWTLLOAPGEAVLVPAGAPHOVQGLVSTISVTQHFLSPETSALSALC 657
Db 1094 LRRRLREMGVSCWTLLOAPGEAVLVPAGAPHOVQGLVSTISVTQHFLSPETSALSALC 1153
QY 658 HQGASLPDPHRLMYAQMDRAVFOAVKAAVAGALQEA 693
Db 1154 HQGASLPDPHRLMYAQMDRAVFOAVKAAVAGALQEA 1189

RESULT 5

US-09-287-354-2
; Sequence 2, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-287-354-2

Query Match 83.4%; Score 3088.5; DB 2; Length 984;
Best Local Similarity 83.0%; Pred. No. 5.9e-282;
Matches 578; Conservative 40; Mismatches 75; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGEVGLTGHKSRRSPLEKQLEEDSSATSEGGGPGPEASLNGK 59
Db 289 LAQCSCQAAGGGGCHACHSQVRRSPGLGELQEBEDTATNSSEEGPGSGDRLSTG 348
QY 60 LAKHLLSGLDRLCLLRLKEREALAWAQREGQGPANTEPGPHCCSRCHGLFNTHWR 119
Db 349 LAKHLLSGLDRLCLLRLKEREALAWAQREGQGPANTEPGPHCCSRCHGLFNTHWR 408
QY 120 CSHCSHRLCVACGRVAGTGRAREKAGFQESAECTQEAAGAACSLMLTQFVSSQALAE 179
Db 409 CPRCSHRLCVACGRVAGTGRAREKAGFQESAECTQEAAGAACSLMLTQFVSSQALAE 468
QY 180 STVHQAWAKFDIRGHCFQVDARVWAPGGGQKQEPTEKTPPTPQSCNGDSNRKDK 239
Db 469 STAHQVWVWFDIRGHCFQVDARVWAPGGGQKQEPTEKTPPTPQSCNGDTHRTKSIK 528
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRTHMAFAPVTPALPSDDRI 299
Db 529 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRTHMAFAPVTPALPSDDRI 588
QY 300 TNLDSIIAQVVERKIQEALGPGLRAGSLRKLGLSLPSVTRLSPPGALLWLQEP - 358

Db 589 TNLDSIIAQVVERKIQEKALGPGLRAGLGLPLSPVPRPLPPPGALLWLQEPQP 648
QY 359 -PKHGFHLFQEHWRQGPVLVSGIQKTLRLSLWMEALGTGGQVQSLTALGPPQPTNLD 417
Db 649 CPRRGHFLFQEHWRQGPVLVSGIQRTLOGNLWGTALGALGGQVQALSPLGPPQSSLG 708
QY 418 STAFWEGFHPETRPKLDGEGSVLLHRTLDGKDSRVQNLVSSLPLPEYCAHOGKUNLAS 477
Db 709 STTFWEGFSPWELRPKSDGEGSVLLHRTLDGKDSRVQNLVSSLPLPEYCAHOGKUNLAS 768
QY 478 YLPLGLTLHLEPQLWAAVGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKD 537
Db 769 YLPPGLALRLEPQLWAAVGVNSHRHGLGTKNLCVEVADLVSLVHARTPLPAWHEAQKD 828
QY 538 FLSGLDGEGLSWPGSGTSTVWHVFRADQRIIRRFLOMVCPCAGAGTLEPGAGSGCYLDAG 597
Db 829 FLSGLDGEGLSWPGSGTSTVWHVFRADQRIIRRFLOMVCPCAGAGTLEPGAGSGCYLDAG 888
QY 598 LRRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSALC 657
Db 889 LRRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSALC 948
QY 658 HQGASLPPDHRMLYAQMDRAVFOAKAAGALQEA 693
Db 949 HOGPSLPPDCHLLYAQMDRAVFOAKAAGALQEA 984

RESULT 6

US-09-287-354-3

; Sequence 3, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-287-354-3

Query Match 83.1%; Score 3077.5; DB 2; Length 1189;
Best Local Similarity 82.9%; Pred. No. 8.6e-281;
Matches 577; Conservative 37; Mismatches 79; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGEVGLTHGSHQSRRLPL-EEKOLEEEDSSATSEEGGGGPGPEASLNGK 59
Db 494 LAQCSCQAAGEGGGHACHSQVRRPLGLGLOEEDTATNSSSEGGSGGPPSRSLTG 553
QY 60 LAKHLLSLGDLRLCLRLKEREALAWAQREGQGPAMTSDSPGPHCCSRCHHGLFNTHWR 119
Db 554 LAKHLLSLGDLRLCLRLGEREALAWAQREGQGPAMTSDSPGPRCCSRCHHGLFNTHWR 613
QY 120 CSHCSHRLVACGRIAGAKNREKTSQEQHTDDCAQEAHAAACSLILTQVSSQALAE 179
Db 614 CPRCSHRLVACGRIAGAKNREKTSQEQHTDDCAQEAHAAACSLILTQVSSQALAE 673
QY 180 STVMHQAWKFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPQSCNGDSNRTKDIK 239
Db 674 STAMHQVWKFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPQSCNGDTHRTKSIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRI 299
Db 734 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRI 793
QY 300 TNLDSIIAQVVERKIQEKALGPGLRAGLGLPLSPVPRPLPPPGALLWLQEPQP - 358

Db 794 TNLDSIIAQVVERKIQEKALGPGLRAGLGLPLSPVPRPLPPPGALLWLQEPQP 853
QY 359 -PKHGFHLFQEHWRQGPVLVSGIQKTLRLSLWMEALGTGGQVQSLTALGPPQPTNLD 417
Db 854 CPRRGHFLFQEHWRQGPVLVSGIQRTLOGNLWGTALGALGGQVQALSPLAPPPQSSLG 913
QY 418 STAFWEGFHPETRPKLDGEGSVLLHRTLDGKDSRVQNLVSSLPLPEYCAHOGKUNLAS 477
Db 914 STTFWEGFSPWELRPKSDGEGSVLLHRTLDGKDSRVQNLVSSLPLPEYCAHOGKUNLAS 973
QY 478 YLPLGLTLHLEPQLWAAVGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKD 537
Db 974 YLPPGLALRLEPQLWAAVGVNSHRHGLGTKNLCVEVADLVSLVHARTPLPAWHEAQKD 1033
QY 538 FLSGLDGEGLSWPGSGTSTVWHVFRADQRIIRRFLOMVCPCAGAGTLEPGAGSGCYLDAG 597
Db 1034 FLSGLDGEGLSWPGSGTSTVWHVFRADQRIIRRFLOMVCPCAGAGTLEPGAGSGCYLDAG 1093
QY 598 LRRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSALC 657
Db 1094 LRRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSALC 1153
QY 658 HQGASLPPDHRMLYAQMDRAVFOAKAAGALQEA 693
Db 1154 HOGPSLPPDCHLLYAQMDRAVFOAKAAGALQEA 1189

RESULT 7

US-09-949-016-7891
; Sequence 7891, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 7891
; LENGTH: 943
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7891

Query Match 63.0%; Score 2333.5; DB 2; Length 943;
Best Local Similarity 77.2%; Pred. No. 9.4e-211;
Matches 438; Conservative 35; Mismatches 63; Indels 31; Gaps 4;
QY 42 TSEGGGGPGPEASLNKGLAKHLLSLGDLRLCLRLKEREALAWAQREGQGPAMTSDSPG 101
Db 387 TAEQAGAGWQEVDRDTSIG-NKQVDSGQD-----EQKGQGPVATEDSPG 429
QY 102 IPHCSCRHCHGLFNTHWRCSHRLCVACRIAGAKNREKTSQEQHTDDCAQEAHAA 161
Db 430 IPHCSCRHCHGLFNTHWRCSHRLCVACRIAGAKNREKTSQEQHTDDCAQEAHAA 489
QY 162 ACSLILTFVSSQALAEALSTVMEHQAQKFDIRGHCFQVDARVWAPGDGQKQKTEPTKTP 221
Db 490 ACSLILTFVSSQALAEALSTVMEHQAQKFDIRGHCFQVDARVWAPGDGQKQKTEPTKTP 549
QY 222 PTPQSCNGDSNRTKDIKTEPTDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDR 281
Db 550 PTPQSCNGDSNRTKDIKTEPTDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDR 609

Db 2315 SPOQHMLSGQPQASHLPQG 2333

RESULT 10

US-09-538-092-1370

; Sequence 1370, Application US/09538092

; Patent No. 6753314

; GENERAL INFORMATION:

; APPLICANT: Giot, Loic

; APPLICANT: Mansfield, Traci A.

; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same

; FILE REFERENCE: 15966-542

; CURRENT APPLICATION NUMBER: US/09/538,092

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: 60/127,352

; PRIOR FILING DATE: 1999-04-01

; PRIOR APPLICATION NUMBER: 60/178,965

; PRIOR FILING DATE: 2000-02-01

; NUMBER OF SEQ ID NOS: 1387

; SOFTWARE: CurapatSeqFormer Version 0.9

; SEQ ID NO 1370

; LENGTH: 2442

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (0)..(0)

; OTHER INFORMATION: Polypeptide Accession Number Q92793

US-09-538-092-1370

Query Match 4.1%; Score 152.5; DB 2; Length 2442;

Best Local Similarity 18.2%; Pred. No. 0.00024;

Matches 156; Conservative 104; Mismatches 282; Indels 317; Gaps 42;

Qy 30 EEKQLEEDSSATSEEGGGPG-----PEASLNKGLAKHLLSLGL 69

Db 1559 EEEERKEEETAASETTESQGSQDSKNAKNNKTKNKNKSKRANKKPKSPMNVNDLS 1618

Qy 70 DRLCLRLKERE-----LAWAQREGQGPAM----- 95

Db 1619 QKLYATMEKHEVFFVIHLHAGPVINTLPVDPDLLSCDLMDGRDAFLTLARDKHWEF 1678

Qy 96 -----TEDSPGPHCCSRCHHGLFNTHWRSCH-SHRLCVAC----- 131

Db 1679 SSLRRSKWSTLCMLVELHTQDRFVYTCNEKHHV-ETRWHTCVCDYDLINCINCYNTKS 1737

Qy 132 -----GRIAGAGNREKTGSOEQHTDCAQEAHAA-----CSLIITQF 170

Db 1738 HAHKWVKGGLDDEGSSQGEPOSKSPQESRRVRSIQRCIOSLVHACOCRNANCSL----- 1792

Qy 171 VSSQALAEISTVMHQAWEKDIRG-----HC-----FC----- 198

Db 1793 PSCQMKRV-VQHTKCKRTNGGCPVKQOLIALCCYHAKGCKENKCPVPFCLNIKHL 1850

Qy 199 ----QVDARVWAPGGGOKE-----PTEKTP-PTPQSPCNGDS 232

Db 1851 RQQQIQHRL-----QQAQLMRRRMATNRNVPOQSLESPTSAPPPTQPS----- 1898

Qy 233 NRTKDIKEETPDSTESPAEDGAGRSPLPCSLCELLASTAVKLCGLGHDRIHMAFPVTPA 292

Db 1899 -----TFQTPQPPAQQPSVSPSPAGFPFVSARTQPTTSTVKPQSQVP---APPPPA 1949

Qy 293 LPSDDRTITLSDISIAQVVERKIQEK-----ALPGCLRAGSLRKLGLSLPLS--- 339

Db 1950 QPPAAVE-----AAQIEREAQQOQLHYRVNINNSMPFG-RFGMTGPSQMAPVSLNV 2002

Qy 340 -----PVRTRLSPGALLWLQEPKXGFLHFOEHWROGQPVVSGIQKTLURLSLWG 391

Db 2003 PRPNQVSGPVMSM-PPCQ--WQAPLIP-----QQQP--MPGLPR-----PVIS 2041

Qy 392 MEALGTGGQVQSITALGPPQ---PTNL-DSTAFWEGFSPHETPK---LDEGSVLLHR 444

Db 2042 MQAQAAG--PRMPSVQPPRSISPSALQDLRLTKGPSPPQOQQOVLNILKSNPQLMAA 2099

Qy 445 TLGDKDASRVQNLVSSSLPLPEYCAHQ-----KLNLSYPLPLGLTLHLEPQLW 493

Db 2100 FIKQRTAKYVANQPGMQPQGLSQPGMQPQPMHQPSLQNLNMAQAGVPRPGVPPQQ 2159

Qy 494 AAYGVN-----SHRGLGTKNLCLVEVSDLI--SILVHAEALPFWRAQKD--- 537

Db 2160 AMGGLNPOQALINMNPNNMNMN--PQYREMLRROLLOOQOQOQOQOQOQOQOQ 2217

Qy 538 ---FLSGLDGEGW---SPGSQTSTVMHVFRQAQRIREF-----QMVCPAG-- 580

Db 2218 SAGMAGMAGHGFQPGPGGYPAM-----QQQRMQOHLPLQSSMGMAAQMGQL 2271

Qy 581 --AGTLEPGAPGSCYLDAGLRRLREEWGVSCTWLLQAPGEAVLVPAGAPHQVQGLVSTI 638

Db 2272 GQMGPGLGADSTPNIQALQORILQQ-----QQMKQOIGSP---GQPNPM 2314

Qy 639 SVTOHFLS--PETSALSAQ 655

Db 2315 SPOQHMLSGQPQASHLPQG 2333

RESULT 11

US-08-227-536-2

; Sequence 2, Application US/08227536

; Patent No. 5658784

; GENERAL INFORMATION:

; APPLICANT: Eckner, Richard

; APPLICANT: Ewen, Mark

; APPLICANT: Livingston, David

; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION

; TITLE OF INVENTION: FACTOR P300 AND USES OF P300

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Weingarten, Schurigin, Gagnebin & Hayes

; STREET: Ten Post Office Square

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/227,536

; FILING DATE: 14-APR-1994

; CLASSIFICATION: 436

; ATTORNEY/AGENT INFORMATION:

; NAME: Williams Ph.D., Kathleen A.

; REGISTRATION NUMBER: 34,380

; REFERENCE/DOCKET NUMBER: DFCI-308XX

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 542-2290

; TELEFAX: (617) 451-0313

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2414 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-227-536-2

Query Match 4.0%; Score 146.5; DB 1; Length 2414;

Best Local Similarity 19.0%; Pred. No. 0.00089;

Matches 141; Conservative 70; Mismatches 242; Indels 289; Gaps 33;

Qy 29 LEE--KQLEEDSSATSEEGGGGPGPE-----ASLNKGLAK--- 62

Db 1513 LEESIKLEEEERKEENTSNESTDVTKGDSKNAKNNKTKNKSLSRGNKKKPG 1572

Qy 63 --HLLSGLDRLCLRLKEREALAWAQREGQPA-----MTEDSPCIP----- 103

```
Db 1573 MPNVSNLQKLYATMEKHEVF-FVIRIAGPAANSPPVDPDPLIPCDLMDGRDAFL 1631
QY 104 -----HCCSRCHHGLFNTHWRSCHC-SHR 126
Db 1632 TLARDKHLFSSLRRAQWSTMCLVELHTQSDRFFVTCNECKHV-ETRWHTVCEDYD 1690
QY 127 LCVACGRIAGAGNREKTG-----SQEHTDDCAQEAAGHAA--- 162
Db 1691 LCITCYNTKNDHDMKELGLGLDDESNNQAAATQSPGDSRRLSIQRCISLVHACQCRN 1750
QY 163 --CSLILTQVSSQALAEISTVHQAWAKFDIGHC-----PCQVD--- 201
Db 1751 ANCSL-----PSCQKMKRV--VQHTGCKRTNGGCPICKQLIALCCYHAKHCQENKCPV 1803
QY 202 -----ARVWAPGDGQOK---BPTKTPPTP---OP 226
Db 1804 PFCNLTKQKLRQOOLQHRLOQAQMLRRRMAWMQRTGVVGGQGLPSPPTPATPTPTGQQP 1863
QY 227 SCNGDSNRTKDIKEETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCGLGHDRIHMAF 286
Db 1864 T-----TPQTPQTSQP-----QTPPNMSM----- 1883
QY 287 APVTPALPDDRTNILDIIAQVVERKIQEAKGGLRAGSLRGLSLPLSPVTRLS 346
Db 1884 ---PPYLP-----RTQAAGP---VSQGAAGQVTPPTPPOTAQ 1916
QY 347 P-PGALLWLQERPP---KHGFHLFQEHWRQGPVLSGQTKTLRLSLWGMEALGTGGQV 402
Db 1917 PLPG-----PPPTAVEMAMQIQRAAETQORAHVQIFORPIQOHMPMTNAPMG--- 1966
QY 403 QSLTALGPPPTNLDSTAFWEGF--SHPETRPKLDGSGVLLHRLTGLDKDASRVQNLVSS 460
Db 1967 -----NNPPMTGRGSGHLEPGMGPTCMQOQPWSSQGL-----PQOQLQSG 2009
QY 461 LPLP---EYCAHQGKLNILASYPLG-LTLHLEPQLWAAAYGVNSHRHGLTKNLCVEVSD 516
Db 2010 MPRPAMMSVAHQOPLNMAPQPLGQVIGISPLKPGTVSQOALQNLRLTRSPSSPLQOQQ 2069
QY 517 LISILVHAEALPPWY---RAOK-----DFLSGLDGLWSPGSGTSTVHVFRAQDAQ 567
Db 2070 VLSIL-HANPQLLAAFIKORAAKYANSNPQIPGQPMGQGLQPTM-----PQOQGV 2124
QY 568 RIRFLQWVCPAGATLEPGAP 589
Db 2125 HSNPAMQNNPMQAGVQVQAGLP 2146
```

RESULT 12

```
US-09-538-092-1289
; Sequence 1289, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqformatter Version 0.9
; SEQ ID NO 1289
; LENGTH: 2414
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q09472
```

US-09-538-092-1289

```
Query Match 4.0%; Score 146.5; DB 2; Length 2414;
Best Local Similarity 19.0%; Pred. No. 0.00089;
Matches 141; Conservative 70; Mismatches 242; Indels 289; Gaps 33;
QY 29 LEE--KQLEEDSSATSEEGGGPGPE-----ASLNKGLAK--- 62
Db 1513 LEESIKELQEBEERKREENTSNESDVTGKDSKNAKKNKKTSKNKSLSRGNKKKPG 1572
QY 63 --HLLSGDLRCLRLRKEREALAWAQREGQPA-----MTEDSGIP--- 103
Db 1573 MPNVSNLQKLYATMEKHEVF-FVIRIAGPAANSPPVDPDPLIPCDLMDGRDAFL 1631
QY 104 -----HCCSRCHHGLFNTHWRSCHC-SHR 126
Db 1632 TLARDKHLFSSLRRAQWSTMCLVELHTQSDRFFVTCNECKHV-ETRWHTVCEDYD 1690
QY 127 LCVACGRIAGAGNREKTG-----SQEHTDDCAQEAAGHAA--- 162
Db 1691 LCITCYNTKNDHDMKELGLGLDDESNNQAAATQSPGDSRRLSIQRCISLVHACQCRN 1750
QY 163 --CSLILTQVSSQALAEISTVHQAWAKFDIGHC-----PCQVD--- 201
Db 1751 ANCSL-----PSCQKMKRV--VQHTGCKRTNGGCPICKQLIALCCYHAKHCQENKCPV 1803
QY 202 -----ARVWAPGDGQOK---BPTKTPPTP---OP 226
Db 1804 PFCNLTKQKLRQOOLQHRLOQAQMLRRRMAWMQRTGVVGGQGLPSPPTPATPTPTGQQP 1863
QY 227 SCNGDSNRTKDIKEETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCGLGHDRIHMAF 286
Db 1864 T-----TPQTPQTSQP-----QTPPNMSM----- 1883
QY 287 APVTPALPDDRTNILDIIAQVVERKIQEAKGGLRAGSLRGLSLPLSPVTRLS 346
Db 1884 ---PPYLP-----RTQAAGP---VSQGAAGQVTPPTPPOTAQ 1916
QY 347 P-PGALLWLQERPP---KHGFHLFQEHWRQGPVLSGQTKTLRLSLWGMEALGTGGQV 402
Db 1917 PLPG-----PPPTAVEMAMQIQRAAETQORAHVQIFORPIQOHMPMTNAPMG--- 1966
QY 403 QSLTALGPPPTNLDSTAFWEGF--SHPETRPKLDGSGVLLHRLTGLDKDASRVQNLVSS 460
Db 1967 -----NNPPMTGRGSGHLEPGMGPTCMQOQPWSSQGL-----PQOQLQSG 2009
QY 461 LPLP---EYCAHQGKLNILASYPLG-LTLHLEPQLWAAAYGVNSHRHGLTKNLCVEVSD 516
Db 2010 MPRPAMMSVAHQOPLNMAPQPLGQVIGISPLKPGTVSQOALQNLRLTRSPSSPLQOQQ 2069
QY 517 LISILVHAEALPPWY---RAOK-----DFLSGLDGLWSPGSGTSTVHVFRAQDAQ 567
Db 2070 VLSIL-HANPQLLAAFIKORAAKYANSNPQIPGQPMGQGLQPTM-----PQOQGV 2124
QY 568 RIRFLQWVCPAGATLEPGAP 589
Db 2125 HSNPAMQNNPMQAGVQVQAGLP 2146
```

RESULT 13

```
PCT-US95-04682-2
; Sequence 2, Application PC/TUS9504682
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurigin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
```

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,536
FILING DATE: 14-April-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holliday C. Heine, Ph.D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-308Xq999
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04682-2

Query Match 4.0%; Score 146.5; DB 4; Length 2414;
Best Local Similarity 19.0%; Pred. No. 0.00089;
Matches 141; Conservative 70; Mismatches 242; Indels 289; Gaps 33;

QY 29 LEB--KLEEDSSATSEGGGPGPE-----ASLNKGLAK--- 62
DB 1513 LEBSEKELEBEERKEENTSNSTVTYKGDKNKKNKNTSKNKSLSRGNKKKPG 1572
QY 63 --HLLSGDLRLCLRLKEREALAWAQEGGPA-----MTEDSPGIP----- 103
DB 1573 MPNVNDSLSQKYATMEKKEVF--FVIRLIAGPAANSLPPIVDPDPLPCDLMDGRDAFL 1631
QY 104 -----HCCSRCHHGLFNTHWRCSHC-SHRLCVAC----- 131
DB 1632 TLARDKHLEFSRLRAQWSTMCLVELHTQSDRFPVYTCNECKHV-ETRWHTCTVCDYD 1690
QY 127 LCVACGRIAGKUREKTG-----SOEQHTDDCAOEAGHAA----- 162
DB 1691 LCITCYNTKXHDHKNKELGLGLDDESNQQAATQSPGDSRRLSIQRCISLVHACORN 1750
QY 163 --CSLILTFVSSQALAEELSTVMHQAQWAKFDIRGHC-----FCQVD----- 201
DB 1751 ANCSL-----PSCQKMKRV--VQHTGCKRKTNGGCPICKQIALCCYHAKHCQENKCPV 1803
QY 202 -----ARWAPGCGQOK-----EPEKTPPTP---OP 226
DB 1804 PFCLNIRKQLRQOQLHRLQQAQMLRRRMAQMORTGVVQGGQLSPPTATPTPTTQGP 1863
QY 227 SCNGDSNRKTDIKETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAF 286
DB 1864 T-----TQQTFTQTSQP-----QTPPNSM----- 1883
QY 287 APVTPALPSDDRIITNILDSTIAQVVERKIQEALGPGLRAGSLRGLSLPLSPVTRLS 346
DB 1884 --PPYLP-----RTQAAGP---VSQKAAGQVTPPTPQTAQ 1916
QY 347 P-PCALLWQPRP---KHGPHLFOEHWRCQOPVLVSGIQKTLRLSLMGWALGTLCQGV 402
DB 1917 PLPG-----PPPTAVEMAMQIQRAAETQMAHVQIFQPIQHPMPPTFMAPMG--- 1966
QY 403 QSLTALGPPQPTNLDSATFMEGF--SHPETRPKLDEGSVLLHRLTGLGDKASRVQNLVSS 460
DB 1967 -----MNPPTMTRGSPGHLEFGMGFTGMQOQPPMSQGL-----PQPPQLOSG 2009

QY 461 LPLP---EYCAHQGKLNLSYLPQG-LTLHPLEPQLMAAYGVNSHRHGLTGKNCVVEUSD 516
DB 2010 MPRPAMMSVAHQGPPLNMAPOGLGVGSPUKPTVTVSOQALQNLLRTLRSFSPLOQQO 2069
QY 517 LISILVHAEALQPPWY---RAQK-----DFLSGLDGEGLWSPSGQSTSTVWHVFRQADAO 567
DB 2070 VLSIL-HANPQLAFAIKORAAKYANSNPQIPGQPGMPQGGPGLQPTM---PQOQGV 2124
QY 568 RIRFLQWVCPAGAGTLEPGAP 589
DB 2125 HSNPAMQNNPMQAGVQORAGLP 2146

RESULT 14
US-08-194-468-2
Sequence 2, Application US/08194468
Patent No. 5750336
GENERAL INFORMATION:
APPLICANT: Montminy, Marc R.
TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN
TITLE OF INVENTION: RESPONSIVE GENES
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,468
FILING DATE: 10-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9672
TELEPHONE: (619)-546-4737
TELEFAX: (619)-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2441 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-194-468-2

Query Match 3.9%; Score 145.5; DB 1; Length 2441;
Best Local Similarity 18.8%; Pred. No. 0.0011;
Matches 158; Conservative 101; Mismatches 290; Indels 301; Gaps 42;

QY 30 EEKQLEEDSSATSEGGGPGP-----PEASLNKGLAKHLJSLGLG 69
DB 1560 EESERKKEESTAASEITEPGSQGDSKNVKNKNTKNKSSISRANKKPSMPNVSNDS 1619
QY 70 DRLCRLLREREAALAWAQEGGQGPATDES PGIP----- 103
DB 1620 QKLYATMEKKEVF-FVIRLIAGPVI STOPPVIDPDPFLSLCDLMDGRDAFLTTLARDKHE 1678
QY 104 -----HCCSRCHHGLFNTHWRCSHC-SHRLCVAC----- 131
DB 1679 FSLRKSMTLCMLVELHTQSDRFPVYTCNECKHV-ETRWHTCTVCDYDLCINCYNK 1737
QY 132 -----GRIAGAKNREKTSQEQHTDDCAOEAGHAA-----CSLILTQ 169
DB 1738 SHTHKVMYKGLGLDDEGSSQGEPSQSPQSRRLSIQRCISLVHACQCRNANCSL----- 1793

This Page Blank (Page 3)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 22:06:54 ; Search time 10 Seconds
(without alignments)
210.046 Million cell updates/sec

Title: US-10-712-629B-18

Perfect score: 3704

Sequence: 1 VTQCSCVQAAGEVGLTGH.....MDRAVFOAVKAAGALQEA 693

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 17545 seqs, 3030971 residues

Total number of hits satisfying chosen parameters: 17545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA_New:*

- 1: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	115	3.1	657	1 US-10-821-234-1477	Sequence 1477, Ap
2	102	2.8	1402	1 US-10-971-982-2	Sequence 2, Appli
3	101	2.7	447	1 US-10-967-527A-14	Sequence 14, Appl
4	97.5	2.6	658	1 US-10-821-234-921	Sequence 921, App
5	97	2.6	614	7 US-11-015-546A-20	Sequence 20, Appl
6	95.5	2.6	1184	1 US-10-131-826A-412	Sequence 412, App
7	95	2.6	4419	1 US-10-821-234-1155	Sequence 1155, Ap
8	93.5	2.5	595	7 US-11-182-946-9	Sequence 9, Appli
9	93	2.5	483	7 US-11-184-574-4	Sequence 4, Appli
10	93	2.5	915	1 US-10-821-234-1514	Sequence 1514, Ap
11	90.5	2.4	542	7 US-11-152-747-6	Sequence 6, Appli
12	90	2.4	777	1 US-10-821-234-1658	Sequence 1658, Ap
13	89.5	2.4	1618	1 US-10-984-645-2	Sequence 2, Appli
14	89	2.4	415	7 US-11-182-946-6	Sequence 6, Appli
15	88.5	2.4	419	1 US-10-821-234-1556	Sequence 1556, Ap
16	88.5	2.4	1138	7 US-11-012-762-28	Sequence 28, Appl
17	88	2.4	1694	7 US-11-135-855-36	Sequence 36, Appl
18	88	2.4	1709	7 US-11-135-855-35	Sequence 35, Appl
19	86.5	2.3	308	1 US-10-967-527A-30	Sequence 30, Appl
20	86.5	2.3	434	7 US-11-135-855-40	Sequence 40, Appl
21	86.5	2.3	653	1 US-10-131-826A-438	Sequence 438, App
22	86.5	2.3	653	7 US-11-135-855-25	Sequence 25, Appl
23	86	2.3	408	7 US-11-135-855-39	Sequence 39, Appl
24	86	2.3	1199	1 US-10-821-234-1126	Sequence 1126, Ap
25	85.5	2.3	1560	7 US-11-059-982-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-10-821-234-1477

; Sequence 1477, Application US/10821234

; Publication No. US20050255114A1

; GENERAL INFORMATION:

; APPLICANT: Labat, Ivan

; APPLICANT: Stache-Crain, Birgit

; APPLICANT: Andarmani, Suean

; APPLICANT: Tang, Y. Tom

; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia

; FILE REFERENCE: 821A

; CURRENT APPLICATION NUMBER: US/10/821,234

; CURRENT FILING DATE: 2004-04-07

; PRIOR APPLICATION NUMBER: US 60/462,047

; PRIOR FILING DATE: 2003-04-07

; NUMBER OF SEQ ID NOS: 1704

; SOFTWARE: pt_seq_genes Version 1.0

; SEQ ID NO 1477

; LENGTH: 667

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-821-234-1477

Query Match 3.1%; Score 115; DB 1; Length 667;

Best Local Similarity 20.4%; Pred. No. 0.046;

Matches 112; Conservative 55; Mismatches 191; Indels 192; Gaps 25;

QY 2 TCQCSCVQAAGE--VGVLT-----GHSQKRRS-----PLEEKQLEEESSA----- 41

Db 60 TKPQDMISAGGVSAGITAIAGKPGDKKXKSLTPAVFVESPDKPSPGSGKGMADALDDL 119

QY 42 -----TSBEGGGGPGPEASLNKGLAKHLGLSLGDLRLRLKREKALAWAQREG-Q 91

Db 120 IDTLGPGPETEENTYTGFEVS--DPMSTVIEELGKREVTPPKYRELL--AKKEGIT 175

QY 92 GPAMTEDSPGIPHCSCRRCHHGLFNTHWRCSHRCLCVACGRAGAKNREKTCGSOEHT 151

Db 176 GPPADSSKPIGPDDAIDALSSDF-----TCGSPTAAGKTKKEESTEVLK 220

QY 152 DCAQBAQHAACSL-----ILTQFVSSQALAEIUSTVM--HQAWAKPDIRGHCFQV 201

Db 221 ---AQSAGTVRSAAAPQEKRKVKEDTMSDQALEALSASLGTRQAPFELDLS--IKEVD 275

QY 202 -----ARYWAPGCGGQ-----KEPTKTPPTPPSCNGCDNRTK 236

Db 276 EKAKEEKEKCEGDEDTIPSEYRLKPATDKGKPLLPPEEKPKPRSESELIDELSEDF 335

QY 237 DIKE--ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPA-- 292

Db 336 DRSECKEKSPKTEKTEESKAAAPAP---VSEAVCRTS--MC-----SIQSAPPBPATL 384
QY 293 ---LPDDRITNILDST-----IAQVVERKIQEKALGPGLRAGSLRKGSLPLSPV 341
Db 385 KGTVP--DDAVEALADSLGKEADPEDGKPVMDKVKEA-----421
QY 342 RTRLSPGALLWLOEPRPKHG-----FHLFOEHWROGQPVLVSGIQKTLRLSLWGM 392
Db 422 -----KEEDREKLGEKETIPDPYRLVEEVKDKGKPLLPKESKEQL-----462
QY 393 EALGTILGGQVQSILALGPPOPTNLDSTAFWEGSHPEPTRPKLDEGSLVLLHRTLGDNDAS 452
Db 463 -----PMSDFLLDALSEDPSGPNASSSLKFEDAKL-----AA 496
QY 453 RVQNLVSSLP 462
Db 497 AISEVVSQTP 506

RESULT 2

US-10-971-982-2
; Sequence 2, Application US/10971982
; Publication No. US20050244889A1
; GENERAL INFORMATION:
; APPLICANT: Rosenfeld, Michael G.
; APPLICANT: Glass, Christopher K.
; APPLICANT: Rose, David W.
; APPLICANT: Torchia, Joseph
; TITLE OF INVENTION: A Transcription Factor Coactivator Protein, p/CIP
; FILE REFERENCE: 6627-PA1021
; CURRENT APPLICATION NUMBER: US/10/971,982
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US/09/445,353
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: PCT/US98/12263
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/049,452
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1402
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312)..(3121)
; OTHER INFORMATION: "n" is any nucleotide
US-10-971-982-2

Query Match 2.8%; Score 102; DB 1; Length 1402;
Best Local Similarity 19.4%; Pred. No. 1.2;
Matches 123; Conservative 75; Mismatches 242; Indels 194; Gaps 27;
QY 22 QKRRSPLEKQLKEEDDSATSEEGGPGPEASLNKGLAKHLLSGLDRLCLLRKERE 81
Db 579 QNSRDPQVKK--ESKSSGEVSETPRGP-LESGHKKLJQLLTCCSDDR-----625
QY 82 ALAWAREGOGPANTEDSRGIHPC-----CSRCHCLFNTHW 118
Db 626 -----GHSLSLTNPLDPNCKDSSVSPSPGSSVSTGTVSTSNVHGSLLQEKH 675
QY 119 RCSHSHRLCVACCRIAGAK-NREKTSQEQHRTDDCAQ-----EAGHAACSL 165
Db 676 R---ILHKLQNGNSPAEVAKITAEATGKTSTASCGEGTTRQEQLSPKKENALLRY 732
QY 166 ILTQVSSQALAEIYSTVHQAWAKFDIRGHCFOVDARVWAPGGGQKQKPTKPTTPQ 225
Db 733 LLDRDDPSDLAK--ELQPQADSGDKLSQCSCSTNP-----SSGQEKDPKIKETNDE 784
QY 226 PSCNGDSNRKTKKEE--TPDSTESP--AEDGAGRSPLCPSPICELLASTAVKLCIGHD 280

Db 785 VS--GDLDNLDAILDGTLSSDFYNNPTNGHGPACAKQOMFAGPSSSLGLRSP-----832
QY 281 RIHMAFAPVTPALPSDDRTITNILDSTIIAQQVVERKIQEKALGPGL-----324
Db 833 -----QPQVSRPPYPNRAVS-LDSPSVSGSPPVKNVSAFPGCLPKQPILAGNPRMWSQ 885
QY 325 -RAGSLRKLGLSLPLSPVTRLSPPGALLWLOEPRPKHGPHLFOEHWROGQPVLVSGIQK 383
Db 886 ENYGAMGNPNRVNPNPT-----SSPG-----DWGLANSRASRMEPLASSPLGR 929
QY 384 TLRLSLWGMALGTL-----GGVQSLTALGPPOPTNLDSTAFWEGSHPEPTRPKLDEGS 438
Db 930 T-----GADYSATLPRPAMGGSVPTL-----PLRSNEL-----PGARPSLQQQQ 968
QY 439 VLLHRTLGDNDASRVQNLVSSLPPEYCAHQKGLNLASYLPLGLTLHLEPQL-----492
Db 969 QQQQQQQQQQQQQQQQQQQQMLQWRTGE-----IPMGVGNVPYPAVQSNQPG 1015
QY 493 -WAAVGVNSHRGHGHTYK---LCVEVSDLSILVHAEQLPPMYRAQK-----FLSLDGL 544
Db 1016 SWPEGMLSMEOGPHGQNRPLLRNSLDLGLPPSNAEGQSD--RALLDQLHTFLSNTDA 1073
QY 545 EG-----LWSPGSGTSTVMHVFRADOA 566
Db 1074 TGLEIDRALGIPELVNVQGALESKQDVFGQEA 1107

RESULT 3

US-10-967-527A-14
; Sequence 14, Application US/10967527A
; Publication No. US20050256041A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Brian A.
; APPLICANT: Holloway, James L.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: Ztafr14, A Tumor Necrosis Factor
; TITLE OF INVENTION: Receptor
; FILE REFERENCE: 03-17
; CURRENT APPLICATION NUMBER: US/10/967,527A
; CURRENT FILING DATE: 2004-10-18
; PRIOR APPLICATION NUMBER: 60/511,698
; PRIOR FILING DATE: 2003-10-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 447
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-967-527A-14

Query Match 2.7%; Score 101; DB 1; Length 447;
Best Local Similarity 23.8%; Pred. No. 0.34;
Matches 72; Conservative 31; Mismatches 119; Indels 80; Gaps 15;
QY 4 CQSCVQAA-GEVGVLT-GHSQKRRSPLEBEKQLEEDSSATSEEGGPGPEASLNKGLA 61
Db 44 CQCPGPGPEEYLSGCGYTK-----DEDYGCVPCEKFKSG--82
QY 62 KHLISGLGRLCLRLKEREALAWAREGOGPANTEDSRGIHCCSRCHHGLFNTHWRCS 121
Db 83 -----GVQICR-RHKDCGFPFRATVLTGPD-MENDAE-----CGPCLFGYYMLENPR 128
QY 122 HCSHRLCVAC-----GRIAGA-----QKNREKTSQEQHRTDDCAQEAAGAACSL 165
Db 129 NIYGMVCSCLLAPPNTKECVATGTSANFPCTSGSSTLSPPQHAHKELSGGCHLATAL 188
QY 166 IL---TQFVSSQALAEIYSTVMH-----QAWAKFDIRGHCFOVDARVWAPGDEGQ--212
Db 189 IIAWSTIFI--MAIAVLIIIMFYILTKPSAPACCTSHPGKSVQAEQVSKDEEKEADPNV 246
QY 213 -----QKEPTEK---TPPTQPCNGDSNRKTKIKETTPDSTESPAEDGAGRSPLPCPSLC 265
Db 247 VMSEKDEFEKLTATPAKPTKSENDASSENEQLLSRSVDSDEEPADPKQG-----SPELC 301


```
QY 266 EL 267
Db 302 LL 303

RESULT 4
US-10-821-234-921
; Sequence 921, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 921
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-921

Query Match 2.6%; Score 97.5; DB 1; Length 658;
Best Local Similarity 22.9%; Pred. No. 1;
Matches 86; Conservative 32; Mismatches 124; Indels 133; Gaps 22;

QY 330 LRKGLSLPLSPVTRLSPGALLMLQEPKHFHLPQEHWRGQPVLSVGIQKTLRLSL 389
Db 21 LEAALGLPVKKRLGRPRGESLRLAE-----VSG----- 50

QY 390 WGMERALGLTGG-VOSTALGPPQPTNLDSATFWEGFSHPETRPKIDEGSVLLHLRLTLD 448
Db 51 -----GGTGLRSALSVPPQPPAGSSRAGSGTG--THTGSDPPMERCA----- 90

QY 449 KDAQRVQNLVSLPLPEYCAHQGKUNLASLPLGLTLHPLEPLQMAAYGVNSHRHGLGTYK 508
Db 91 -GAGRKLPTGRCPVTE-----GSTVQLIAP--WNAADVHSH--GDK 127

QY 509 N-LCVEVS---DLISLVHAEALPPWYRAQKDFLSGLDGEGLW-----SGSQTS 555
Db 128 DSQTCIRVSASDPRLPKKEEELAPLLRTHLQAE-----PHQGCWTVTPEAAMTPGNATP 183

QY 556 TVWHVFRQAQDAQIRRFQMVCPAGAGTL---EPGAPGSCYLDAGLRRRLREEMGV--- 608
Db 184 P-----RTEVTPRLRLQL-LPLANTTLSTPNPDQASAPDP---RPLREEBEARLL 234

QY 609 -----SCWTLQAQGEAVLVPGA-PHQVQGLVSTISVTQHFLSPETSALSAQ 655
Db 235 PRTHLQAEHLHQGCWTVTE---PAALTGNATPRTQEVTPLELLEQL--PE----- 282

QY 656 LCHQASLSP-PDHRM 669
Db 283 LVHATLSTPNPNQV 297

RESULT 5
US-11-015-546A-20
; Sequence 20, Application US/11015546A
; Publication No. US20050250126A1
; GENERAL INFORMATION:
; APPLICANT: GAO, ZEREN
; APPLICANT: SHEPPARD, PAUL O.
; APPLICANT: FOX, BRIAN A.
; APPLICANT: HOLLOWAY, JAMES L.
; APPLICANT: JASPERS, STEPHEN R.
; APPLICANT: APPELBY, MARK

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24

; TITLE OF INVENTION: ztnf13, A TUMOR NECROSIS FACTOR
; FILE REFERENCE: 03-24
; CURRENT APPLICATION NUMBER: US/11/015,546A
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: 60/530,185
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial DNA sequence, polypeptide sequence of
; OTHER INFORMATION: the MBP-ztnf13 fusion
US-11-015-546A-20

Query Match 2.6%; Score 97; DB 7; Length 614;
Best Local Similarity 22.8%; Pred. No. 1.1;
Matches 79; Conservative 27; Mismatches 138; Indels 102; Gaps 16;

QY 54 ASLNKGLAKHLLSG--LGDRLCELLRKEREALAWAORE-----GQGPAMT-----E 97
Db 270 ASPNKLAKBFLFNLYLLTDEGLEAVNKKOKPLGAVALKSYEEELAKDPRIATATMENAKQGE 329

QY 98 DSPGIPHCSCRCHHGLFNTHWRCSCSHRLCVACGRI-AGAGKNREKTSQEQHTDDCAQ 156
Db 330 IMPNIPQ-----MSAFW-----YAVRTAVINAASGRQTVDEALKDAQTSSSH 372

QY 157 EAGHAACSLILTFVSSQALABELSTMHQAWAKFIRGHCFCQVDARVWAPGGGQOKEP 216
Db 373 HHHHHANSVPLVP-----RGSREFGTAGSAMIPGGATARAEP 409

QY 217 TEKTPPTQPSGNGSDNRTKDIKEETPD-----STESPAEDGAG-----RSLPLCPSLCELL 268
Db 410 AR--PRAGAPRTRGSGR-----EPFGGAAATTTTPPGGAGSARAPTPAPPAPAP 461

QY 269 ASTAVKLCIGHDRIRHMAFAPVTPA---LPSDDRITNILDSTIAQVVERKIOEKALGPGL 324
Db 462 A-----AYAPVGAATALELPASGIGHVQTT-----VLRGGADGRAAAAL 502

QY 325 RAGSGLRKGL-----SLPLSPVTRLSPPGALLMLQEPKHFHFL 365
Db 503 RGAHGHARLPQDRHALPESP--RQRGEAGAAASQLQAALPGPLHL 547

RESULT 6
US-10-131-826A-412
; Sequence 412, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
```


; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-182-946-9

Query Match 2.5%; Score 93.5; DB 7; Length 595;
Best Local Similarity 20.6%; Pred. No. 1.9;
Matches 119; Conservative 65; Mismatches 186; Indels 209; Gaps 35;
QY 1 VTQCQSC-----VOAAGVGVLTGHSOKSR-----RSPLEEKOLEBEDSSATS 43
DB 119 VNSCARCFHSPAGMIVKFPFGTAQNTVCEPASPGVSPACASPENCK-----EPSSGTI 174
QY 44 BEGGGGPGPSPASLNKGLAKHLLSGLGDLRCR-----LLRKEBALAWAQRGQGPAMTEDS 99
DB 175 PQ--AKPTVPSPATSSASTMPVRG--GTRLAQEAASKLTRAPDSFSSVGRPSSDPGLSPTQ 231
QY 100 P---GIPHCSSRCHHGLF--NTHWRCSCSHRLCVACGRIAGAGKREKTSQEQHTDDCA 155
DB 232 PCPEGSDCKQCEPDYILDEAGRCT-----ACVSCSR-----DDL 268
QY 156 QEAGHAACSLITQFVSSQALBELSTMHQAQAKFDIRGHC-----FCQVDARYVWPGDGG 211
DB 269 EK---TPCAW-----NSSRTCECRPGMICATSNATSCARCVYPICAAET--VTKPDMA 318
QY 212 QQKEPTEKPTPTQPSGNGSDNRKDKIETPTSTSPAB-----DGGRSPLPCP 262
DB 319 EKDTTPEAPLGTQPCDN-----PTPENGAPASTPTQSLLVDSQASKTLP 367
QY 263 S-----LCELLASTAVKLCGLGHR-----IHWAPA 287
DB 368 TSAPVALSSTGKVPDLGAPVFWLVLVGVVSSAPFLLC--HRRACRKRIRKQJHLCY 424
QY 288 PVTPLPSPDDRITNILDIIIAQVVERKIQEALGPGLRAGSGRLKGLSL--PLSPVTRL 345
DB 425 PVQTSQPK-----LVLDSRPR-----RSSTQLRSGASVTPEVAERGLM 464
QY 346 SPP-----GAL-----LWQEPKPHG-----FHLFOEHWRRQGPVLVSGIQKT 384
DB 465 SQPLMETCHSVGAAYLESPLQDASPAGGSPSRDLPEPRVSTEHTN-----NKIEK 516
QY 385 LRLSLWGMEA-----LCTLGQVSLTAL--GPPQP-----TNLSTAFWEGFSHPETPKL 434
DB 517 ----IYIMKADTVIVGTVXAELEPEGRGLAGPAPSELEEELEADHTPH---YPEQETEPPL 569
QY 435 DEGS-VLLHLHRTLGDKDASRVQNLVSSLPPEYCAHQK 472
DB 570 GSCSDVWLSVEEGKED-----PLP--TAASGK 595

RESULT 9
US-11-184-574-4
; Sequence 4, Application US/11184574
; Publication No. US2005025541A1
; GENERAL INFORMATION:
; APPLICANT: Meng, Xiang-Jin
; APPLICANT: Haqshenas, Gholamreza
; APPLICANT: Huang, Fang-Fang
; TITLE OF INVENTION: AVIAN HEPATITIS E VIRUS, VACCINES AND METHODS OF PROTECTING
; TITLE OF INVENTION: AGAINST AVIAN HEPATITIS-SPLENOMEGALY SYNDROME AND MAMMALIAN
; TITLE OF INVENTION: HEPATITIS E
; FILE REFERENCE: AM100389
; CURRENT APPLICATION NUMBER: US/11/184,574
; CURRENT FILING DATE: 2005-07-19
; PRIOR APPLICATION NUMBER: US/10/029,840
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 483

; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-11-184-574-4

Query Match 2.5%; Score 93; DB 7; Length 483;
Best Local Similarity 22.7%; Pred. No. 1.6;
Matches 88; Conservative 39; Mismatches 120; Indels 140; Gaps 24;
QY 285 AFAPVTPALPSPDDRITNILDIIIAQV--VERKIQEALGPGLRAGSGRLKGLSLPLSPVTR 343
DB 6 AAPAGVLRAPDDTVQGLLDIPPAHTDVAALVTAETAIIG-----HAPLE--LA 50
QY 344 RLSPPGALLWLOPRPKHGFHLFQEHWRQG---QPVLVSGIQKTLRLSLMG----- 391
DB 51 AINPPGPVL-----EQGLLYMPARLDGRDEWVKQLQSLDTHVHCLRAAP 92
QY 392 ---MEALGTLGGOVQSLTALGPPQPTNLDSTA--FWE--GFSPHET-----RPKLD 435
DB 93 TSLAVINTLVGRYGRATKLPVEYDLMMDTIAQFHHHIGPINFPSTLEYAEMCEAMLSKGQ 152
QY 436 EGSVLLHLHRTLGDKDASRV---QNLVSSLPPEYCAHQKLNLSYLPGLTLHPLRPQL 492
DB 153 DGS-LIVHLDLQADCSRITFFQKCAKFTLDDPVAH--GKVG-----QGISAMP----- 199
QY 493 WAAVGVNSHRHGLGTGNLCVEVSDLSILVHBAQALPPWYRA--QKDFLSGLDGBGLWSPG 551
DB 200 -----KTLIC-----ALFGPWFAIEKHLVAGL-----PPG 224
QY 552 SQTSTVHVFPRAQDAQIRRFLOMVC--PAGAGTLEPGAQ-----SCYLDAGLRRRL 602
DB 225 YVYG---DLTYEADLHR-----SVLCAPAGHLVFENDSFEDSTQNNVSLDLECELRRF 276
QY 603 -REEWGV-----SCWTLLOAPGEAV 621
DB 277 GMPDMVVALYHLVRSYV--LLVAPKEAL 302

RESULT 10
US-10-821-234-1514
; Sequence 1514, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 1514
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1514

Query Match 2.5%; Score 93; DB 1; Length 915;
Best Local Similarity 21.8%; Pred. No. 3.6;
Matches 157; Conservative 61; Mismatches 257; Indels 246; Gaps 37;
QY 54 ASLNKGLAKHLLSGLGDLRLCRLLRKEREALAWAQRGQGPAMTEDSPGIPHCSSRCHHGL 113
DB 7 AGLDEGALRKLLEVTD-----LAERRRISAIRELQCELEEREALASKRFAERQD 60
QY 114 FNTHWRCSCSHRLC---VACGRIAGAGKREKTSQEQHTDDCAQEAQ-----HAA 162
DB 61 NKENW--LHSQQREARQAALARLAG---QLESMDNVEELT--ALLRSAGEYBERKLIRAA 114
QY 163 CSLILTQFVSSQALA--ELSTVMHQWAKFDIRG---HCF--CQVDARVWAPGGGQKPE 216

```

Db      115 IRRVRAQEI EAATUAGRLYSGRPNSSGRSDSKGLAAHRLQCEVPER-----EQSQQ 167
Qy      217 TEKTPPTPOPSCNGSDNRKDIKEET-----PDSTESPAEDAGRSPLPCPSCCELLAS 270
Db      168 AEVSKPIPTP-----EGTSQDVTTVTLRLRAPPGSTSS--SPASPSSSPTPASPEPPLE 219
Qy      271 TAVKLCGLGHRDHMAFAPVTP-ALPSSDRITN-----ILDSIIAQVVERKIQKALGP 322
Db      220 PAEAQCL-----TAEVPGSEPPSPPKTTSPEPQESFTLPSTEGVQVNNKLLSGPKETP 273
Qy      323 GLRAGSGLRKGLSLPLSPVRTRLSPPGALLMQEPRPKHGPHLQFEHWRQGPVLVSGIQ 382
Db      274 AAQ-----SPRPGSDTKRAD-----VAGPRPCQ-----297
Qy      383 KTLRLSLWGMERALGTGLGQVQSLTALGPQPT-NLDSTAFMEGFGSHPETRPKLDGSLVLL 441
Db      298 -----RSLSVLSPRQPAQNRESTPLASG--PSSFOR--AGSV--330
Qy      442 LHRTLGDKDSRVONLVSSLPPLPEYCAHOGKLNLASYLPLG---LTLHLEPOLWAAVGV 498
Db      331 -----RDRVHKFTSDSP-----MAARLQDGTPOAALSPLTTPR--KAPGP 367
Qy      499 NSHRGHLGTKNLCVEVSDLSILVHAEALQPEWYRAQKDFLSGLDGEGLWSPGSOTSTVM 558
Db      368 LPHQH-----PCVLLQRLLSG-----PSTSSRPS 394
Qy      559 HVFR--AQDAQIRIRFLQMVCPAGAG-----TLEPGAPGSCYLDAGLRR 600
Db      395 KEQGVQAQPLAQLRSCPOEGPRGRGLAARPLENRAGGPVARSEPGAPLPVAVGTA---451
Qy      601 RLREWGVSCTLL-----QAPGRAVLVPAG-----APHQV---QGLVS 636
Db      452 ----EPGDSMKTTFTIEIKDGRGQASTGRVLLPTGNQRAELTLGLRAPPTLLSTSGGKS 507
Qy      637 TISVTHFLSPETSALSQALCH--QGASLPPDHR-----MLYAQMDRAVFOAKAAGVAL 689
Db      508 TITRVN---SPGTLARLGSVTHVTSFSHAPSSRGCSIKWEAEPAEPLAAVEAANGAE 564
Qy      690 Q 690
Db      565 Q 565

RESULT 11
US-11-152-747-6
; Sequence 6, Application US/11152747
; Publication No. US20050251881A1
; GENERAL INFORMATION:
; APPLICANT: E. I . du Pont de Nemours, Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Luan
; TITLE OF INVENTION: CAROTENOID KETOLASE GENE
; FILE REFERENCE: CL-1849 US NA
; CURRENT APPLICATION NUMBER: US/11/152,747
; CURRENT FILING DATE: 2005-06-14
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 542
; TYPE: PRF
; ORGANISM: Synecocystis sp. PCC6803
US-11-152-747-6

```

```

Qy 589 PGSCYLDAGLRRRLREEWGVSCTWLLQAPGEAVILVAGAPHQVQGLVSTI-----SVTQ 642
Db 147 PPQALLDLAL-----NYG---WENL-----KSVLAITAGSKTKALDFIRTNIGSPEDVINE 193
Qy 643 HFLSPETSALSALQCHQ-GASLPPDHR-----MLYAQMDFRAVFQAKAAVGAALQEA 692
Db 194 WFDSEVKAPLARLCSIGA--PPSQGSSSGMMVMAMRHLEGIARPKGTGALTEA 248

RESULT 12
US-10-821-234-1658
; Sequence 1658, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: dt_seq_genes Version 1.0
; SEQ ID NO 1658
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1658

```

```

Query Match      2.4%; Score 90.5; DB 7; Length 542;
Best Local Similarity 27.1%; Pred.No.2.9;
Matches 48; Conservative 20; Mismatches 66; Indels 43; Gaps 10;

538 FLSGLDGEGLWSPGSTTWWHV--PRADAQRIIRF-----LQWVCPAGAGTLEPCA 588
|||||
93 FCGGLGQAPMSYRSLEKTCAHATYSPDAEKYQFVNYWTDDLNAVOPAF-----NA 146
|||||

```

```
QY 614 LOAPGBAVLPAGAPHOVQGLVS-----TISVTQHFLSPETSALSAQ 655
Db 535 -----EVLGSVGP---TFLVSCDRPSTGGDEAPTPAPLLRLAQHKMKWPSVSSLDSA 585
QY 656 L-----CHQASLPPDH 667
Db 586 LESSPSLHSPAD--PSH 600

RESULT 13
US-10-984-645-2
; Sequence 2, Application US/10984645
; Publication No. US20050244386A1
; GENERAL INFORMATION:
; APPLICANT: Habener, Joel
; APPLICANT: Zulewski, Hendrik
; APPLICANT: Abraham, Elizabeth
; APPLICANT: Vallejo, Mario
; TITLE OF INVENTION: METHOD OF TRANSPLANTING IN A MAMMAL AND TREATING DIABETES MELLITU
; TITLE OF INVENTION: BY ADMINISTERING A PSEUDO-ISLET LIKE AGGREGATE DIFFERENTIATED FR
; TITLE OF INVENTION: A NESTIN-POSITIVE PANCREATIC STEM CELL
; FILE REFERENCE: 3284/1223
; CURRENT APPLICATION NUMBER: US/10/984,645
; CURRENT FILING DATE: 2004-11-09
; PRIOR APPLICATION NUMBER: US 09/731,255
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 60/169,082
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/215,109
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: US 60/239,880
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-984-645-2

Query Match 2.4%; Score 89.5; DB 1; Length 1618;
Best Local Similarity 23.5%; Pred. No. 14;
Matches 79; Conservative 23; Mismatches 121; Indels 113; Gaps 18;

QY 8 VQAAGSVGLTGHOSKRSPLKEEEDSATSSEGGGPGPEASLNKGLAKHLGSG 67
Db 1105 VGLGDPGHLT-----REEVNPEPLEEESLEAKVQGLEP-----RKDLER--AGG 1149
QY 68 LGDRLCRL-----RKERALAWAQREGQGPAMT----- 96
Db 1150 LGTEFSELPGKSRDPWEPREGRESEAEAPGAEAFPAETLHTGSDAPSPWPLGSEE 1209
QY 97 -----ESPGI-PHCCSRCHHGLFNTHWRCSHCHSLRCVACGR1A 135
Db 1210 ABEEDVPVLVSPSTYTPILEDAPGLQPOA-----EGSQEASWGVQ-----GR-A 1253
QY 136 GAGK---NREKTSQSOHTDDCAQAGHAACSILTLQFVSSQALAE-----LST 181
Db 1254 EAKVSEBQELGSGE--IPEGLQEEGES-----RESEEDLGTLPDSTPLGYLRS 1306
QY 182 VMHQAWAKFDIRGHCFQVDARVWAPGDGQQ--KEPTEKTP-PTPQSCNGDSNRKTDI 238
Db 1307 PTSRPTWPLESRGHPLKGTKEGMDPAVLASEGLEPSEKERESEEGEECGRDS-----DL 1362
QY 239 KEETPD-STESPAEDGA-GRSPLPCPSLCELLASTA 272
Db 1363 SEEFEDLGTEAPFLPGVGEVAEPPLQGVPLQLLDPA 1398

RESULT 14
US-11-182-946-6
; Sequence 6, Application US/11182946
; Publication No. US20050255100A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Wei, Ying-Fei
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
; FILE REFERENCE: 1488.1280004
; CURRENT APPLICATION NUMBER: US/11/182,946
; CURRENT FILING DATE: 2005-07-18
; PRIOR APPLICATION NUMBER: US/10/186,643
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US/09/573,986
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-182-946-6

Query Match 2.4%; Score 89; DB 7; Length 415;
Best Local Similarity 19.9%; Pred. No. 2.7;
Matches 76; Conservative 51; Mismatches 171; Indels 84; Gaps 18;

QY 106 CSRCHHGLFNTHWRCSHCSH-RLCVACGRIAGAGKNREKTGSOEQHTDDCAQAGHAACS 164
Db 80 CKTCPHNSYNEHW--NHLSTCQLCRPCDIVLGFEEVAPCTSDRKA---ECRCQPG-MSCV 133
QY 165 LILTOPVSSQALAEISTVMHQAWAKFDIRGHCF-CQVDARVWAPGDGQOQKEPTEKTPPT 223
Db 134 YLDNECVHCE---EERLVLCQPGTEAEVTDMDTDCVPCPKPGHFQNTSSPRARCQPH 190
QY 224 PQPSCNGDSNRKTDIKEETPDST-----ESPAEDGA-----GRSPLPC- 261
Db 191 TRCEIQG-----LVAAAPGTSYSDTICKNPPPGAMLLALLSLVLFLFTTTLACA 243
QY 262 ----PSLCELLASTAVKLCGLGHDRIHMAFAPVTPALPSPDDRITNLDLSIIAQVVERKIQE 317
Db 244 WNRHPSLCKRLGLTLLKRHPGEGE-----SPCPAPRADPHDPDLAEPLLP--MSGDLS 295
QY 318 KALGPGLRAGSLRKLGLSLPLSP-----VTRLSPGALLMLQEPKPKGHFLFQBHWRO 372
Db 296 SPAGP-----PTAPSLAEVVLQQSP---LVQARELEAEPEGHQVAVGAN 338
QY 373 GPVLVSGTQKTLRLSLWGMELGTLGGQVQSLTALGPPQPTNLDSTAFEGFSHPE--T 430
Db 339 GIHVTGGSVTVTGNIIYINGPVLGTRGPGDPAPPEPPYPTP-----EEGAPGSPSELS 392
QY 431 RPKLDEGSV--LLLHRTLGDKD 450
Db 393 TPYQEDGKAWHLAETETLGCQD 414

RESULT 15
US-10-821-234-1556
; Sequence 1556, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1556
; LENGTH: 419
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1556

Query Match      2.4%; Score 88.5; DB 1; Length 419;
Best Local Similarity 21.1%; Pred No. 3;
Matches 92; Conservative 45; Mismatches 172; Indels 127; Gaps 21;

QY 67 GLGDRLCRLRKEREALAWAQREGQGFAMTEDSPGIP-----HC---CSR 108
Db 32 GUGER-C-YSEAQRCDGSDCADGTD---EEDCGCPGPHPCGAAGTSGATACYLPADR 86
QY 109 CHHGLF-----NTHWRCSHSHRLCVACGRITAGAKNREKTSQE-----QHTDDCAOEAGH 160
Db 87 CNYQTECADGADERRCRHCQ-----PGNFRCRDEKCVYETWVCDGQPCDADGSD 136
QY 161 AACSLJLTQFVSSOAL-----AELSTVMHQAWAKFDIRGHCFQVDAR 203
Db 137 WDCSYVLPRKVIITAAVIGSLVCGLLLVIALGCTCKLYAIRTQYSIFAPLSRMEAEIVQQ 196
QY 204 VWAPGDG---GQOKEPTKPTTPQPSGNGDSNRTKDI-----KEETPDSTESPAEDGAG 255
Db 197 QAPPSYGQLIAQGAIPVEDPFTENPDNSVLGNLSLLQILRQDMTPGGGPGARRRQRG 256
QY 256 RSPLPSPCLCELLASTAVKLCGHDRIHMAFAPVTPALPSDDRITNILDIIIAQVVERKI 315
Db 257 R-----LMRRLVRRL-----RRWGLLPRTNTPARASEAR-----SQVTPSAA 293
QY 316 QEKAL-----GPGLRAGS-GLRKGLSLPLSPVTRL-----SPPGALLWLQEPKPKGHFLF 366
Db 294 PLEALDGGTGPAREGGAVGGQGEQAPPLFIKAPLPSASTSPATTVPPEAGPLPSLPL- 352
QY 367 QEHWQGPVLVSGIQKTLRLSLWGMALGTGLGQVQSQTALGPPQFTNLDSTAFWEGFS 426
Db 353 -----EPSLLSGVVQALR-----CRL-----LPSLGGPPGTRSPPG-----P 384
QY 427 HPETRPKLDGSGVILL 442
Db 385 HTAVLALEDEDDVLLV 400
```

Search completed: November 25, 2005, 22:17:23
Job time : 14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 22:00:09 ; Search time 232 Seconds
(without alignments)
2107.462 Million cell updates/sec

Title: US-10-712-629B-18

Perfect score: 3704

Sequence: 1 VTQCSCVQAAGEVGLTGH.....MDRAVFOAVKAAVGAALQEAQ 693

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3682	99.4	1182	1	HAIR MOUSE
2	3663.5	98.9	1181	2	Q4QY90_MOUSE
3	3510	94.8	1181	1	HAIR RAT
4	3122.5	84.3	1187	1	Q8WNL9_MACMU
5	3095.5	83.6	1189	1	HAIR HUMAN
6	3095.5	83.6	1189	2	Q6G830_HUMAN
7	1534	41.4	342	2	Q863E6_PIG
8	840.5	22.7	709	2	Q8H276_SHEEP
9	729	19.7	1265	2	Q9P2G7_HUMAN
10	728.5	19.7	2081	2	Q5S028_HUMAN
11	728.5	19.7	2162	2	Q8N3U0_HUMAN
12	719.5	19.4	1441	2	Q5RCM5_PONPY
13	719	19.4	2210	2	Q723E7_HUMAN
14	702	19.0	2428	2	Q692K6_MOUSE
15	700.5	18.9	1417	2	Q9NYF4_HUMAN
16	700.5	18.9	1578	2	Q9BW93_HUMAN
17	700.5	18.9	1761	2	Q7LBCE_HUMAN
18	700.5	18.9	1787	2	Q9UP80_HUMAN
19	679.5	18.3	989	2	Q6P9S0_MOUSE
20	653	17.6	1331	2	Q61RB8_XENLA
21	641.5	17.3	1325	2	Q521X8_CHICK
22	636.5	17.2	1658	2	Q4T2W0_TETNG
23	634	17.1	759	2	Q9BVH6_HUMAN
24	617	16.7	1334	2	Q5HZN1_XENLA
25	610.5	16.5	793	2	Q5U5V7_MOUSE
26	608.5	16.4	744	2	Q8K2J6_MOUSE
27	608.5	16.4	749	2	Q8K2K4_MOUSE
28	608.5	16.4	1323	2	Q6PCM1_MOUSE
29	608.5	16.4	1334	2	Q6ZQ57_MOUSE
30	606.5	16.4	1321	2	Q68D47_HUMAN
31	606.5	16.4	1321	2	Q53872_HUMAN

32	606.5	16.4	1338	2	Q9Y4C1_HUMAN	Q9Y4C1 homo sapien
33	604.5	16.3	1214	1	TSGA_RAT	Q63679 rattus norv
34	603.5	16.3	1212	2	Q9P0K0_HUMAN	Q9P0K0 homo sapien
35	602.5	16.3	1325	2	Q6U0T9_HUMAN	Q6U0T9 homo sapien
36	598.5	16.2	1321	2	Q6N0S0_HUMAN	Q6N0S0 homo sapien
37	550.5	14.9	854	2	Q9VHCS_DROME	Q9VHCS drosophila
38	549.5	14.8	710	2	Q7QCCL1_ANOGA	Q7QCCL1 anopheles g
39	539	14.6	955	2	Q5SQZ9_HUMAN	Q5SQZ9 homo sapien
40	493.5	13.3	492	2	Q8K2A5_MOUSE	Q8K2A5 mus musculu
41	490.5	13.2	444	2	Q8CCB2_MOUSE	Q8CCB2 mus musculu
42	471.5	12.7	592	2	Q8R350_MOUSE	Q8R350 mus musculu
43	458.5	12.4	359	2	Q8C5Q9_MOUSE	Q8C5Q9 mus musculu
44	448.5	12.1	1192	2	Q8IY08_HUMAN	Q8IY08 homo sapien
45	366.5	9.9	968	2	Q6NV48_MOUSE	Q6NV48 mus musculu

ALIGNMENTS

RESULT 1

HAIR_MOUSE
ID HAIR_MOUSE STANDARD; PRT; 1182 AA.
AC Q61645; Q80Y47;
DT 01-NOV-1997 (Rel. 35, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hairless protein.
GN Name=Hr;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BALB/c; TISSUE=Skin;
RX MEDLINE=94329587; PubMed=8052649;
RA Cachon-Gonzalez M.B., Fenner S., Coffin J.M., Moran C., Best S.,
RA Stoye J.P.;
RT "Structure and expression of the hairless gene of mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:7717-7721(1994).
[2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP STRAIN=C57BL/6; TISSUE=Retina;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May act as a transcription factor that could act on to
regulate one of the phases of hair growth.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Expressed predominantly in brain, hair
follicles and interfollicular epidermis. No expression in dermis.
CC -!- DISEASE: Defects in Hr are the cause of a number of pleiotropic
effects including structural abnormalities of epithelial cells in
the hair follicles, hair loss towards the end of the first hair
CC

CC growth cycle, and the failure of subsequent hair growth cycles.
CC Older mice carrying an hr mutation have been reported to possess
CC altered ratios of T-cell-dependent B-cell responses. Mice
CC homozygous for hr mutation are uniquely sensitive to UV and
CC chemically induced skin tumors.
CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----

DR ENBL; Z22675; CAA83587.1; -; mRNA.
DR ENBL; BC049182; AAH49182.1; -; mRNA.
DR PIR; I48378; I48378.
DR Ensembl; ENSMUSG0000022096; Mus musculus.

DR MGI; MGI:96223; Hr.
DR InterPro; IPR003347; TF_JmJc.

DR Pfam; PF02373; JmJc; 1.

DR SMART; SM00558; JmJc; 1.

KW DNA-binding; Metal-binding; Nuclear protein; Transcription;

KW Transcription regulation; Zinc; Zinc-finger.

FT ZN FING 595 620 C6-type.

FT COMPTAS 535 540 Poly-Gly.

FT CONFLICT 401 401 P -> S (in Ref. 1).

FT SEQUENCE 1182 AA; 127193 MW; 3APABE96C6EB3241 CRC64;

Query Match 99.4%; Score 3682; DB 1; Length 1182;

Best Local Similarity 99.4%; Pred. No. 2.1e-235;

Matches 689; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VTQCQCVQAAGEVGLTGHQSOKSRSPLEEKQLEEDSSATSEEGGGGPGPEASLNKGL 60

DB 490 VTQCQCVQAAGEVGLTGHQSOKSRSPLEEKQLEEDSSATSEEGGGGPGPEASLNKGL 549

QY 61 AKHLLSGLDRLCLRLKEREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 120

DB 550 AKHLLSGLDRLCLRLKEREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 609

QY 121 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 180

DB 610 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 669

QY 181 TVMHOAWAKFDIRGHCFQVDARVWAPGQGOQKEPTEKTPPTPPQSCGDSNRTKDIKE 240

DB 670 TVMHOAWAKFDIRGHCFQVDARVWAPGQGOQKEPTEKTPPTPPQSCGDSNRTKDIKE 729

QY 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300

DB 730 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 789

QY 301 NILDSIIAQVVERKIQEALGPGLRAGSLRGLSLPLSPVTRLSPPGALLWLQEPRPK 360

DB 790 NILDSIIAQVVERKIQEALGPGLRAGSLRGLSLPLSPVTRLSPPGALLWLQEPRPK 849

QY 361 HGPHLFOEHWRCQOPVLVSGIQKTLRLSLWGMALGTLGGQVQTLTALGPPQPTNLDSTA 420

DB 850 HGPHLFOEHWRCQOPVLVSGIQKTLRLSLWGMALGTLGGQVQTLTALGPPQPTNLDSTA 909

QY 421 FWEGFHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSIPLPEYCAHOGKLNLSYLP 480

DB 910 FWEGFHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSIPLPEYCAHOGKLNLSYLP 969

QY 481 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLS 540

DB 970 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLS 1029

QY 541 GLDGEGLWSFGSQTSTVWHVFPRAQDAQRIIRFLQWVCPAGAGTLEPAGPSCYLDAGLR 600

DB 1030 GLDGEGLWSFGSQTSTVWHVFPRAQDAQRIIRFLQWVCPAGAGTLEPAGPSCYLDAGLR 1089

QY 601 RLREEMGVCWTLLOAPGEAVLVPAGAPHQVGLVSTISVTOHFLSPETSALSQALCHQG 660

DB 1099 RLREEMGVCWTLLOAPGEAVLVPAGAPHQVGLVSTISVTOHFLSPETSALSQALCHQG 1149

DB 1090 RLREEMGVCWTLLOAPGEAVLVPAGAPHQVGLVSTISVTOHFLSPETSALSQALYHQG 1149

QY 661 ASLPPDHRMLYAOQMDRAVFPQAVKAAVAGALQEA 693

DB 1150 ASLPPDHRMLYAOQMDRAVFPQAVKAAVAGALQEA 1182

RESULT 2

Q4QY90 MOUSE

ID Q4QY90 MOUSE PRELIMINARY; PRT; 1181 AA.

AC Q4QY90;

DT 13-SEP-2005 (trEMBLrel. 31, Created)

DT 13-SEP-2005 (trEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (trEMBLrel. 31, Last annotation update)

DE Hairless protein.

GN Name=Hr; Synonyms=hr;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=KM; TISSUE=Skin;

RX PubMed=15955095; DOI=10.1111/j.0022-202X.2005.23744.x;

RA Zhang J.T., Fang S.G., Wang C.Y.;

RT "A Novel Nonsense Mutation and Polymorphisms in the Mouse Hairless

Gene.";

RL J. Invest. Dermatol. 124:1200-1205(2005).

DR EMBL; AY547391; AAT45233.1; -; mRNA.

DR MGI; MGI:96223; Hr.

SQ SEQUENCE 1181 AA; 127068 MW; 53DEF4A87D8207BE CRC64;

Query Match 98.9%; Score 3663.5; DB 2; Length 1181;

Best Local Similarity 99.1%; Pred. No. 3.6e-234;

Matches 687; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 VTQCQCVQAAGEVGLTGHQSOKSRSPLEEKQLEEDSSATSEEGGGGPGPEASLNKGL 60

DB 490 VTQCQCVQAAGEVGLTGHQSOKSRSPLEEKQLEEDSSATSEEGGGGPGPEASLNKGL 548

QY 61 AKHLLSGLDRLCLRLKEREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 120

DB 549 AKHLLSGLDRLCLRLKEREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 608

QY 121 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 180

DB 609 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 668

QY 181 TVMHOAWAKFDIRGHCFQVDARVWAPGQGOQKEPTEKTPPTPPQSCGDSNRTKDIKE 240

DB 669 TVMHOAWAKFDIRGHCFQVDARVWAPGQGOQKEPTEKTPPTPPQSCGDSNRTKDIKE 728

QY 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300

DB 729 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 788

QY 301 NILDSIIAQVVERKIQEALGPGLRAGSLRGLSLPLSPVTRLSPPGALLWLQEPRPK 360

DB 789 NILDSIIAQVVERKIQEALGPGLRAGSLRGLSLPLSPVTRLSPPGALLWLQEPRPK 848

QY 361 HGPHLFOEHWRCQOPVLVSGIQKTLRLSLWGMALGTLGGQVQTLTALGPPQPTNLDSTA 420

DB 849 HGPHLFOEHWRCQOPVLVSGIQKTLRLSLWGMALGTLGGQVQTLTALGPPQPTNLDSTA 908

QY 421 FWEGFHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSIPLPEYCAHOGKLNLSYLP 480

DB 909 FWEGFHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSIPLPEYCAHOGKLNLSYLP 968

QY 481 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLS 540

DB 969 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLS 1028

QY 541 GLDGEGLWSPGSGTSTVMHVFRQAQRIIRRFLOWMPCPAGAGTLEPGAGSGCYLDAGLRR 600
|||||
Db 1029 GLDGEGLWSPGSGTSTVMHVFRQAQRIIRRFLOWMPCPAGAGTLEPGAGSGCYLDAGLRR 1088
|||||
QY 601 RUREEWSVSWTLLQAPGEAVLPAGAPHQVGLVSTISVTOHFLSPETSALSQAQLCHOG 660
|||||
Db 1089 RUREEWSVSWTLLQAPGEAVLPAGAPHQVGLVSTISVTOHFLSPETSALSQAQLCHOG 1148
|||||
QY 661 ASLPPDHRLMYAQMDRAVFAVKAAGALQAEK 693
|||||
Db 1149 ASLPPDHRLMYAQMDRAVFAVKAAGALQAEK 1181
|||||

RESULT 3

HAIR_RAT
ID HAIR_RAT STANDARD; PRT; 1181 AA.
AC P97609;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hairless protein.
GN Name=Hr;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=97141510; PubMed=8987811;
RA Thompson C.C.;
RT "Thyroid hormone-responsive genes in developing cerebellum include a
novel synaptotagmin and a hairless homolog.";
RL J. Neurosci. 16:7832-7840(1996).
CC -!- FUNCTION: May act as a transcription factor that could act on to
regulate one of the phases of hair growth.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
DR EMBL; U71293; AAC53018.1; ALT INIT; mRNA.
DR Ensembl; ENSRNOG0000011427; Rattus norvegicus.
DR InterPro; IPR003347; TF_JmJc.
DR Pfam; PF02373; JmJc; 1.
DR SMART; SM00558; JmJc; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
FT ZN FING 594 619 C6-type.
SQ SEQUENCE 1181 AA; 127307 MW; 834B7029CFB588F0 CRC64;

Query Match 94.8%; Score 3510; DB 1; Length 1181;
Best Local Similarity 95.1%; Pred. No. 5.5e-224;
Matches 659; Conservative 11; Mismatches 21; Indels 2; Gaps 2;

QY 1 VTQCSQVQAAGEVGLTGHSGKRRSPLEEKQLEEDSSATSEBGGGPGPEASLNKGL 60
|||||
Db 491 VTQCSQVQAAGEVGLTGHSGKRRSPLEEKQLEEDSSATSEBGGGPGPEASLNKGL 548
|||||
QY 61 AKHLLSGLGDLRLCLLRKEREALAWAOREGQGPANTEDESGIPHCSCRHGLFNTHWR 120
|||||
Db 549 AKHLLSGLGDLRLCLLRKEREALAWAOREGQGPANTEDESGIPHCSCRHGLFNTHWR 608
|||||
QY 121 SHCSHRLCVACGRIAGAGNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 180
|||||
Db 609 SHCSHRLCVACGRIAGAGNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 668
|||||
QY 181 TVMHQWAKFDIRGHCFQVDARVWAPDGGQOKEPTEKTPPTQPSGNSNRKTDIKE 240
|||||

Db 669 TVMHQWAKFDIRGHCFQVDARVWAPDGGQOKEPTEKTPPTQPSGNSNRKTDIKE 728
|||||
QY 241 ETPDSTESPAERAGSRPLPCPSLCELLASTAVKLCGHDRIHMAFAPVTPALPSDDRIT 300
|||||
Db 729 ETPDSTESPAERAGSRPLPCPSLCELLASTAVKLCGHDRIHMAFAPVTPALPSDDRIT 788
|||||
QY 301 NILDSIIAQQVVERKIQEAKLGPCLRAGSGLRKGLSLPLSPVTRTLSPGPGALLWQPRPK 360
|||||
Db 789 NILDSIIAQQVVERKIQEAKLGPCLRAGSGLRKGLSLPLSPVTRTLSPGPGALLWQPRPK 848
|||||
QY 361 HGFHLFQEHWRQGPVLVSGIOKTLRLSLWGEALGTLCGGQVOSLTALGPQPTNLDSTA 420
|||||
Db 849 HGFHLFQEHWRQGPVLVSGIOKTLRLSLWGEALGTLCGGQVOTLTALGPQPTSLDSTA 908
|||||
QY 421 FWEFGSHPETRPKLDGSGVLLHRTLTGDKASRVQNLVSSLPLPEYCAHOGKUNLASLYP 480
|||||
Db 909 FWKGFSPHPEARPKLDGSGVLLHRLPLGDKDGRVENLASSLPLPEYCAHOGKUNLASLYP 968
|||||
QY 481 LGTLHLPLEPQWAAVGVNSHRHGLTGNLCVVEVSDLSILVHAEALQPPWYRAQKDFLS 540
|||||
Db 969 LGTLHLPLEPQWAAVGVNSHRHGLTGNLCVVEVSDLSILVHAEALQPPWYRAQKDFLS 1028
|||||
QY 541 GLDGEGLWSPGSGTSTVMHVFRQAQRIIRRFLOWMPCPAGAGTLEPGAGSGCYLDAGLRR 600
|||||
Db 1029 GLDGEGLWSPGSGTSTVMHVFRQAQRIIRRFLOWMPCPAGAGTLEPGAGSGCYLDAGLRR 1088
|||||
QY 601 RUREEWSVSWTLLQAPGEAVLPAGAPHQVGLVSTISVTOHFLSPETSALSQAQLCHOG 660
|||||
Db 1089 RUREEWSVSWTLLQAPGEAVLPAGAPHQVGLVSTISVTOHFLSPETSALSQAQLCHOG 1148
|||||
QY 661 ASLPPDHRLMYAQMDRAVFAVKAAGALQAEK 693
|||||
Db 1149 ASLPPDHRLMYAQMDRAVFAVKAAGALQAEK 1181
|||||

RESULT 4

Q8WNL9 MACMU
ID Q8WNL9 MACMU PRELIMINARY; PRT; 1187 AA.
AC Q8WNL9;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Hairless.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopitheidae; Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ahmad W., Ratterre M.S., Panteleyev A.A., Aita V.M., Sundberg J.P.,
RA Christiano A.M.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF361864; AAL56245.1; -, mRNA.
DR InterPro; IPR003347; TF_JmJc.
DR Pfam; PF02373; JmJc; 1.
DR SMART; SM00558; JmJc; 1.
SQ SEQUENCE 1187 AA; 127134 MW; D1CFED4AAA76C77E CRC64;

Query Match 84.3%; Score 3122.5; DB 2; Length 1187;
Best Local Similarity 83.6%; Pred. No. 2.8e-198;
Matches 582; Conservative 43; Mismatches 68; Indels 3; Gaps 2;
QY 1 VTQCSQVQAAGEVGLTGHSGKRRSPLEEKQLEEDSSA-TSEBGGGPGPEASLNKG 59
|||||
Db 492 LAQCSQAQAAGEGGGPGAGHFOQVRRSPGLGSLQEEEDTAANSSSEEGSGPGDGLSTG 551
|||||
QY 60 LAKHLLSGLGDLRLCLLRKEREALAWAOREGQGPANTEDESGIPHCSCRHGLFNTHWR 119
|||||
Db 552 LSKHLLSGLGDLRLCLLRKEREALAWAOREGQGPANTEDESGIPHCSCRHGLFNTHWR 611
|||||
QY 120 CSKSHRLCVACGRIAGAGNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 179
|||||

Db 612 CPCSHRLCVACGRVAGAGRAREKAGSREQSTECTQEAHAAACSLTLTFQVSSQALAE 671
QY 180 STVHWQAWKFDIRGHCFQOVDARVWAPGQGOKEPTKTPPTPQSCNGDSNRKTDIK 239
Db 672 STAHQVWVFDIRGHCPQOVDARVWAPGQGOKESTQKTPPTPQSCNGDTHRYSIK 731
QY 240 EETPDSTESPAEDAGAGSPRLPCPSLCELLASTAVKJCLGHDRHMAFAPVTPALPSSDRI 299
Db 732 EETPDSTETPAEDRAGQAPLPCPSLCELLASTAVKJCLGHERHMAFAPVTPALPSSDRI 791
QY 300 TNLDSIIAQVVERKIQEAKGLGRLAGSLPLSPVTRLSPPGALLWLOEPRP 359
Db 792 TNLDSIIAQVVERKIQEAKGLGRLAGSLPLSPVTRLSPPGALLWLOEPRP 851
QY 360 K--HGPHLFOEHWRGQPVLSVGTQKTLRLSLMGMEALGTLGQVOSLTLGALPQPPTNLD 417
Db 852 RPORGHFLFOEHWRGQPVLSVGIQRTLQNLGWTALGALGQVQALSPLGPPQPTSLG 911
QY 418 STAFWEGFHPETRPKLDGESSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKUNLAS 477
Db 912 STAFWEGFHPETRPKLDGESSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKUNLAS 971
QY 478 YLPLGLTLHLEPOLWAAVGVNSHRGLGKTLNLCVEVDLSILVHAEAOPLPWYRAQKD 537
Db 972 YLPLGLTLHLEPOLWAAVGVNSHRGLGKTLNLCVEVDLSILVHAEAOPLPWYRAQKD 1031
QY 538 FLSLGDEGLWSPGSGTSTVHVFRQAQDQIRRFQWCPAGAGTLEPGAPGSCYLDAG 597
Db 1032 FLSLGDEGLWSPGSGTSTVHVFRQAQDQIRRFQWCPAGAGTLEPGAPGSCYLDAG 1091
QY 598 LRRLREWGVSWTLLQAPGAEAVLPAGAPHQVQGLVSTISVTQHFSPETSALSALQC 657
Db 1092 LRRLREWGVSWTLLQAPGAEAVLPAGAPHQVQGLVSTISVTQHFSPETSALSALQC 1151
QY 658 HQGASLPDPHRLMYAQNDRVAFQAVKAVGALQEA 693
Db 1152 HQGASLPDPHRLMYAQNDRVAFQAVKAVGALQEA 1187
RESULT 5
HAIR_HUMAN STANDARD; PRT; 1189 AA.
AC O43593; Q96H33; Q9NPE1;
DT 15-JUL-1999 (Rel. 38, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hairless protein.
GN Name=HR;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND VARIANT ALUNC ALA-1022.
RX MEDLINE=9411413; PubMed=9445480; DOI=10.1126/science.279.5351.720;
RA Ahmad W., ul Haque M.F., Brancolini V., Tsou H.C., Ul Haque S.,
RA Lam H., Aita V.M., Owen J., Deblaquiere M., Frank J.,
RA Cserhalmi-Friedman P.B., Leake A., McGrath J.A., Peacocke M.,
RA Ahmad M., Ott J., Cristiano A.M.;
RT "Alopecia universalis associated with a mutation in the human hairless gene.";
RL Science 279:720-724 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), SEQUENCE REVISION TO 572 AND 774, AND TISSUE SPECIFICITY.
RC Tissue=Peripheral blood leukocyte, and skin fibroblast;
RX MEDLINE=99162400; PubMed=10051399; DOI=10.1006/geno.1998.5699;
RA Ahmad W., Zlotogorski A., Panteleyev A.A., Lam H., Ahmad M.,
RA ul Haque M.F., Abdallah H.M., Dragan L., Cristiano A.M.;
RT "Genomic organization of the human hairless gene (HR) and identification of a mutation underlying congenital atrichia in an Arab Palestinian family.";
RL Science 279:720-724 (1998).

Genomics 56:141-148(1999).
[3]
RN SEQUENCE REVISION TO 446 AND 584.
RP Cristiano A.M.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE (ISOFORMS 1 AND 2), VARIANT ALUNC ASP-1136, AND TISSUE SPECIFICITY.
RC Tissue=Brain, Fetal brain, and peripheral blood leukocyte;
RX MEDLINE=98409496; PubMed=9736769; DOI=10.1093/hmg/7.11.1671;
RA Cichon S., Anker M., Vogt I.R., Rohleder H., Putzstuck M., Hillmer A.,
RA Farooq S.A., Al-Dhafri K.S., Ahmad M., Haque S., Rietschel M.,
RA Propping P., Kruse R., Noethen M.M.;
RT "Cloning, genomic organization, alternative transcripts and mutational analysis of the gene responsible for autosomal recessive universal congenital alopecia.";
RL Hum. Mol. Genet. 7:1671-1679(1998).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 198-1189 (ISOFORM 3).
RC Tissue=Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP VARIANT GLN-620.
RX MEDLINE=98431781; PubMed=9758627;
RA Ahmad W., Irvine A.D., Lam H., Buckley C., Bingham E.A.,
RA Panteleyev A.A., Ahmad M., McGrath J.A., Cristiano A.M.;
RT "A missense mutation in the zinc-finger domain of the human hairless gene underlies congenital atrichia in a family of Irish travellers.";
RL Am. J. Hum. Genet. 63:984-991(1998).
RN [7]
RP VARIANT GLN-620.
RX MEDLINE=21303039; PubMed=11410842;
RA Hillmer A.M., Kruse R., Betz R.C., Schumacher J., Heyn U.,
RA Propping P., Noethen M.M., Cichon S.;
RT "Variant 1859G-->A (Arg620Gln) of the 'hairless' gene: absence of RT association with papular atrichia or androgenic alopecia.";
RL Am. J. Hum. Genet. 69:235-237(2001).
RN [8]
RP VARIANT ALUNC ASN-1012.
RX MEDLINE=22294554; PubMed=12406339;
RA DOI=10.1046/j.1523-1747.2002.00268.x;
RA Klein I., Bergman R., Indelman M., Sprecher E.;
RT "A novel missense mutation affecting the human hairless thyroid receptor interacting domain 2 causes congenital atrichia.";
RL J. Invest. Dermatol. 119:920-922(2002).
RN [9]
RP -1- FUNCTION: May act as a transcription factor that could act on to regulate one of the phases of hair growth.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=Long;
CC IsoId=O43593-1; Sequence=Displayed;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]_TaxID=9606;
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RX MEDLINE=22198257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RA Strausberg R.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC067128; AA067128.1; -; mRNA.
 SQ SEQUENCE 1189 AA; 127494 MW; 67A4B95A01063387 CRC64;

Query Match 83.6%; Score 3095.5; DB 2; Length 1189;
 Best Local Similarity 83.2%; Pred. No. 1.8e-96;
 Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;

QY 1 VTCQSCVOAAGEVGLTGHSSKRSPL-EEQLEEDSSATSEEGCGGPGPEASINLKG 59
 DB 494 LAQCSCQAAGGEGGHACHSQVRRSPGLGELQOEEDTAINSSSEEGPGGDSRUSTG 553
 QY 60 LAKHLLSGLGDLRLRLKREALAWAQREGQGPAMTEDSPGPHCCSRCHGLFNTHWR 119
 DB 554 LAKHLLSGLGDLRLRLKREALAWAQREGQGPAMTEDSPGPHCCSRCHGLFNTHWR 613
 QY 120 CSHCSHRLCVACGRIAGAKNREKTSQBOHTDDCAQBAHACSLILTOFVSSQALAE 179
 DB 614 CPRCSHRLCVACGRIAGAKNREKTSQBOHTDDCAQBAHACSLILTOFVSSQALAE 673
 QY 180 STVMHQAWKFDTRGHCFQVDARVWAPGGGQKQKTEPTPTPPQSCNGDSNRTKDIK 239
 DB 674 STAHQVWVFXFDTRGHCFQVDARVWAPGGGQKQKTEPTPTPPQSCNGDTHRTKSIK 733
 QY 240 EETPDSTESPAEDGAGRSPLPCPSLCBLASTAVKLCIGHDR IHMAFAPVTPALPSDDRI 299
 DB 734 EETPDSTESPAEDGAGRSPLPCPSLCBLASTAVKLCIGHDR IHMAFAPVTPALPSDDRI 793
 QY 300 TNILDSIIAQVVERKI QEKALGPGLRAGSGRLKGLSLPLSPVTRTLSPPGALLWQEP- 358
 DB 794 TNILDSIIAQVVERKI QEKALGPGLRAGSGRLKGLSLPLSPVTRTLSPPGALLWQEP 853
 QY 359 -PKHGFHLFOHWHQOGQVPLVSGTQKTLRLSLWCMEALGTGGQVQSITALGPPQPTNLD 417
 DB 854 CPRRGFHLFOHWHQOGQVPLVSGTQKTLRLSLWCMEALGTGGQVQSITALGPPQPTNLD 913
 QY 418 STAFWEGFSHPETPKLDEGSVLLHRTLGDPKDASRVQNVLSSPLPEYCAHQKLNLAS 477
 DB 914 STTFWEGFSHPETPKLDEGSVLLHRTLGDPKDASRVQNVLSSPLPEYCAHQKLNLAS 973
 QY 478 YLPLGLTLHPLQPLWAAVGNVSHRHLGTQKLVCEVSDLSILVHAEALPPWYRAQKD 537

RESULT 7

Q863E6 PIG PRELIMINARY; PRT; 342 AA.
 AC Q863E6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hairless (Pigment).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
 OC Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22755563; PubMed=12873232;
 RX DOI=10.1046/j.1365-2052.2003.01032.x;
 RA Fernandez A., Silio L., Noguera J.L., Sanchez A., Ovilo C.;
 RT "Linkage mapping of the porcine hairless gene (HR) to chromosome
 RT 14.";
 RL Anim. Genet. 34:317-318(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Fernandez A.I., Silio L., Rodriguez C., Noguera J.L., Sanchez A.,
 RA Ovilo C.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY279972; AAP33389.1; -; mRNA.
 DR Interpro: IPR003347; TF_JmJC.
 DR Pfam: PF02373; JmJC; 1.
 DR SMART: SM00558; JmJC; 1.
 FT NON_TER 1 1
 FT TER 342 342
 SQ SEQUENCE 342 AA; 37147 MW; 35F00261F07B4A42 CRC64;

Query Match 41.4%; Score 1534; DB 2; Length 342;
 Best Local Similarity 84.8%; Pred. No. 1.5e-93;
 Matches 285; Conservative 15; Mismatches 36; Indels 0; Gaps 0;

QY 329 GLRKGSLPLSPVTRTLSPGALLWLQEPKPHGLFQSHWRQGPVLVSGTQKTLRLS 388
 DB 2 GLRKGSLPLSPVTRTLSPGALLWLQEPKPHGLFQSHWRQGPVLVSGTQKTLRLS 61
 QY 389 LWCMEALGTGGQVQSITALGPPQPTNLDSTAFWEGFSHPETPKLDEGSVLLHRTLGD 448
 DB 62 LWETEALGALGGQVHALTPLGPPQSTSLGSTAFWEGFSRPTETPKSDEGSVLLHRTLGD 121
 QY 449 KDSARVQNVLSSPLPEYCAHQKLNLASYLPLGLTLHPLQPLWAAVGNVSHRHLGK 508
 DB 122 EDASRVENLAASPLPEYCAHKGKLNLASYLPPGGLRPLVPOIWAAYGVSHPHGLGK 181
 QY 509 NLCVEVSDLSILVHAEALPPWYRAQKDFLSGLDGEGLWSGSGTSTVWHVFAQAQOR 568
 DB 182 NLCVEVTSVLVHAEALPFTWHRQAQDFLSGLDGEGLWSGSGTSTVWHVFAQAQOR 241
 QY 569 IRRFLOWVCPAGAGTLEPGAGSGCYLDAGLRRLRREEWGVCWTLTQAPGEAVLVPAGAP 628
 DB 242 IRRFLOWVCPAGAGTLEPGAGSGCYLDAGLRRLRREEWGVCWTLTQAPGEAVLVPAGAP 301

```
QY 629 HQVQGLVSTISVTHFLSPETSALSALCHQASLP 664
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 302 HQVQGLVSTISVTHFLSPETSALSALCHQASLP 337

RESULT 8
Q8H276 SHEEP PRELIMINARY; PRT; 709 AA.
AC Q8H276;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hairless (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Finocchiaro R., Portolano B., Damiani G., Caroli A., Budelli E.,
RA Bella P., Pagnacco G.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY130969; AA05753.1; -, mRNA.
FT NON TER 709
SQ SEQUENCE 709 AA; 75435 MW; 838EC14E3A1B0071 CRC64;

Query Match 22.7%; Score 840.5; DB 2; Length 709;
Best Local Similarity 70.1%; Pred. No. 3.9e-47;
Matches 155; Conservative 16; Mismatches 49; Indels 1; Gaps 1;

QY 1 VTCQCVOAAGVGVLTHGSKRSPLLEKLEEDSSATSEGGGPGPEASLNKGL 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 490 ITQCQCAAGTGTGPAHYSQPLSLLLGEGPQDEDSAAASKGEGGS-GPEAGLSVGL 548

QY 61 AKHLLSGDLRLCLRLKREALAWAQREGQGPAMTEDSPGIPHCSCRHGHLFNTHWC 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 549 AKHLLSALGDLRLCLRLKREALAWAQREGQAPVGTEDDPSLPRCCSCHCHGLFNTHWC 608

QY 121 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQEAAGHAAACSLILTFVSSQALAE 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 609 PRCSHRLCVTCGRMAGASTREKAGSGQSTKCGQAGHAGSLMLTFVSSQALAE 668

QY 181 TVMHQAWAKFDIRGHCFCDVADRVAPAGDGGQKEPTEKTP 221
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 669 AMHQVWVAFDIRGHCFCDQADRVWTAGDGLQKEPTEKTP 709

RESULT 9
Q9P2G7 HUMAN PRELIMINARY; PRT; 1265 AA.
AC Q9P2G7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE KIAA1380 protein (Fragment).
GN Name-KIAA1380;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirotsawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
DR EMBL; AB037801; BAA92618.1; -, mRNA.
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OX Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Howden P.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL713895; CAIL10948.1; -; Genomic DNA.
DR EMBL; AL590502; CAIL10948.1; JOINED; Genomic DNA.
DR EMBL; AL590502; CAIL10948.1; JOINED; Genomic DNA.
DR Ensemble; ENSG00000171988; Homo sapiens.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003347; TF_JmJC.
DR Pfam; PF02373; JmJC; 1.
DR SMART; SM00558; JmJC; 1.
KW Receptor.
KW NCBI_TaxID=9606;
SQ SEQUENCE 2081 AA; 231066 MW; 4DED8856C50CEDCE CRC64;
Query Match 19.7%; Score 728.5; DB 2; Length 2081;
Best Local Similarity 29.1%; Pred. No. 3.8e-39;
Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;
QY 61 AKHLISGLDRLCRLLRERALEAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 120
DB 1343 SKYLDIIGDKFCQVLTSEKTSALSWVKDAK-IWKRAVRGVREMCDAEATLNFNIHWVC 1401
QY 121 SHCSHRLCVACGRIAGAGKREKTSQEQHT-DDCAQEAAGHAACSLILTFVSSQALAE 179
DB 1402 QKGFVVCCLDYK-----AKERKSRDKELYAMWKVKQGHDKHLMPTQIIIGSVLTDL 1457
QY 180 STVMHQAWAKFDIRGHCF-----QVDARVWAPGDGGQ-----KEPTE 218
DB 1458 LDAMHTLREKYGIKSHCHCTNKQNLQVGNPPTWNGVSQVLQNVLNHNSKISLCWPSQOQ 1517
QY 219 KTPPTPOPSGNGSNRTKDIEE---TPDSTESP----- 249
DB 1518 LDAMHTLREKYGIKSHCHCTNKQNLQVGNPPTWNGVSQVLQNVLNHNSKISLCWPSQOQ 1575
QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCGLGHDRIHMAFAPV--- 289
DB 1576 QIKEREQDNSESPNGRTSPLVSNQNEQGSTLRDLTTAGKLRVSGTDAGIAFAPVYSM 1635
QY 290 -TPALPDDDRITNILDIIIAQVVERKI-----QEKALPGRLRAGSGLR 331
DB 1636 GAPSSKSGRTMPNILDIIASVVENKIPPSKTSKINVKPELKEEPEESIIISAVDENNKLY 1695
QY 332 KGLSLPLSPVTRTLSPGALLWLQEPKPGHFLFOEHWKQGPVLVSGIQKTLRLSLWG 391
DB 1696 S--DIPHSWICEK-----HILWLKDYKNSSNWKLFKECWKQGPVAVSGVHKMKNISLWK 1748
QY 392 MEALG-TLGGQVQSALTALGPPQPTNLDSTAFWEGFHPETRPKLDGESSVLLHRLTLDGD 450
DB 1749 AESISLDFGHDQADLLCKDSIIISNANVKFWDGFEVSKRQKNKSETVVLK-----KD 1804
QY 451 -----ASRVQNLVSSLPLPEYCAHQKUNLASYPGLGTLHP-LEPOLWAAYGV 498
DB 1805 WPSGEDEKTMMPARYEDLLKSLPLPEYCNPEKFNASHLP-GFFVRPDLGPRLCISAYG 1863
QY 499 NSHRGH-LGTNLCVEVSDILISILVHAEAQLPWPYRAQKDFLSGLDGEGLWS----- 549
DB 1864 VAAKDHDIGTITNLHIEVSDVNVILVYVGIAGKNGILSKAGILKKFEEDDDIILRKLDK 1923
QY 550 ----PGSQTSTVMHVFRQAQRIIRFLQWVC-PAGAGTL---BFGAPGSCYLDAGLRRLR 601
DB 1924 SSEIPGA---LWHIYAGKVDKIREFLQKISKEQGLEVLPEHPDIRDQSVYVKNLQR 1979
QY 602 LREWGVSCWTLTLLQAPQEAULVPAGAPHQVQGLVSTISVTOHFLSPETSALSQALCHQGA 661
DB 1980 LLEBYGVRTCTLIQFLGDIAIVLPAGALHQVQNFHSCIQVTEDFVSPHLVESFHLTQELR 2039
QY 662 SLPP-----DHRMLYAQMDRAVFOAKAAGVALQ 690
DB 2040 LLKBEINYDDKL---QVKNILYHAVKENVRLAK 2069
```

```
RESULT 11
QY 61 AKHLISGLDRLCRLLRERALEAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 120
DB 1424 SKYLDIIGDKFCQVLTSEKTSALSWVKDAK-IWKRAVRGVREMCDAEATLNFNIHWVC 1482
QY 121 SHCSHRLCVACGRIAGAGKREKTSQEQHT-DDCAQEAAGHAACSLILTFVSSQALAE 179
DB 1483 QKGFVVCCLDYK-----AKERKSRDKELYAMWKVKQGHDKHLMPTQIIIGSVLTDL 1538
QY 180 STVMHQAWAKFDIRGHCF-----QVDARVWAPGDGGQ-----KEPTE 218
DB 1539 LDAMHTLREKYGIKSHCHCTNKQNLQVGNPPTWNGVSQVLQNVLNHNSKISLCWPSQOQ 1598
QY 219 KTPPTPOPSGNGSNRTKDIEE---TPDSTESP----- 249
DB 1599 NTPPKSEK--NGSSSPESDVGTNKLTPPESQSPHLWADLABQKAREEKENKELTLEN 1656
QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCGLGHDRIHMAFAPV--- 289
DB 1657 QIKEREQDNSESPNGRTSPLVSNQNEQGSTLRDLTTAGKLRVSGTDAGIAFAPVYSM 1716
QY 290 -TPALPDDDRITNILDIIIAQVVERKI-----QEKALPGRLRAGSGLR 331
DB 1717 GAPSSKSGRTMPNILDIIASVVENKIPPSKTSKINVKPELKEEPEESIIISAVDENNKLY 1776
QY 332 KGLSLPLSPVTRTLSPGALLWLQEPKPGHFLFOEHWKQGPVLVSGIQKTLRLSLWG 391
DB 1777 S--DIPHSWICEK-----HILWLKDYKNSSNWKLFKECWKQGPVAVSGVHKMKNISLWK 1829
QY 392 MEALG-TLGGQVQSALTALGPPQPTNLDSTAFWEGFHPETRPKLDGESSVLLHRLTLDGD 450
DB 1830 AESISLDFGHDQADLLCKDSIIISNANVKFWDGFEVSKRQKNKSETVVLK-----KD 1885
QY 451 -----ASRVQNLVSSLPLPEYCAHQKUNLASYPGLGTLHP-LEPOLWAAYGV 498
DB 1886 WPSGEDEKTMMPARYEDLLKSLPLPEYCNPEKFNASHLP-GFFVRPDLGPRLCISAYG 1944
QY 499 NSHRGH-LGTNLCVEVSDILISILVHAEAQLPWPYRAQKDFLSGLDGEGLWS----- 549
DB 1945 VAAKDHDIGTITNLHIEVSDVNVILVYVGIAGKNGILSKAGILKKFEEDDDIILRKLDK 2004
```

```
Query Match 19.7%; Score 728.5; DB 2; Length 2162;
Best Local Similarity 29.1%; Pred. No. 4e-39;
Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;
QY 61 AKHLISGLDRLCRLLRERALEAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 120
DB 1424 SKYLDIIGDKFCQVLTSEKTSALSWVKDAK-IWKRAVRGVREMCDAEATLNFNIHWVC 1482
QY 121 SHCSHRLCVACGRIAGAGKREKTSQEQHT-DDCAQEAAGHAACSLILTFVSSQALAE 179
DB 1483 QKGFVVCCLDYK-----AKERKSRDKELYAMWKVKQGHDKHLMPTQIIIGSVLTDL 1538
QY 180 STVMHQAWAKFDIRGHCF-----QVDARVWAPGDGGQ-----KEPTE 218
DB 1539 LDAMHTLREKYGIKSHCHCTNKQNLQVGNPPTWNGVSQVLQNVLNHNSKISLCWPSQOQ 1598
QY 219 KTPPTPOPSGNGSNRTKDIEE---TPDSTESP----- 249
DB 1599 NTPPKSEK--NGSSSPESDVGTNKLTPPESQSPHLWADLABQKAREEKENKELTLEN 1656
QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCGLGHDRIHMAFAPV--- 289
DB 1657 QIKEREQDNSESPNGRTSPLVSNQNEQGSTLRDLTTAGKLRVSGTDAGIAFAPVYSM 1716
QY 290 -TPALPDDDRITNILDIIIAQVVERKI-----QEKALPGRLRAGSGLR 331
DB 1717 GAPSSKSGRTMPNILDIIASVVENKIPPSKTSKINVKPELKEEPEESIIISAVDENNKLY 1776
QY 332 KGLSLPLSPVTRTLSPGALLWLQEPKPGHFLFOEHWKQGPVLVSGIQKTLRLSLWG 391
DB 1777 S--DIPHSWICEK-----HILWLKDYKNSSNWKLFKECWKQGPVAVSGVHKMKNISLWK 1829
QY 392 MEALG-TLGGQVQSALTALGPPQPTNLDSTAFWEGFHPETRPKLDGESSVLLHRLTLDGD 450
DB 1830 AESISLDFGHDQADLLCKDSIIISNANVKFWDGFEVSKRQKNKSETVVLK-----KD 1885
QY 451 -----ASRVQNLVSSLPLPEYCAHQKUNLASYPGLGTLHP-LEPOLWAAYGV 498
DB 1886 WPSGEDEKTMMPARYEDLLKSLPLPEYCNPEKFNASHLP-GFFVRPDLGPRLCISAYG 1944
QY 499 NSHRGH-LGTNLCVEVSDILISILVHAEAQLPWPYRAQKDFLSGLDGEGLWS----- 549
DB 1945 VAAKDHDIGTITNLHIEVSDVNVILVYVGIAGKNGILSKAGILKKFEEDDDIILRKLDK 2004
```

550 QY 550 -----PGSQTSTVHVHFRADQAOIRRRFLQVMVC-PAGAGTL---EPGAPGSCYLDAGLRRR 601
 2005 DB 2005 SSEIPGA----LWHIYAGKVDVKIREFLQKISKEQGLEVLPEHDPIRDQSWYVNNKLRQ 2060
 602 QY 602 LREEMGVSWTLLQAPGEAVLPACAPHOVQGLVSTISVTOHFLSPETSALSQAOLCHQGA 661
 2061 DB 2061 LBEYGVRTCTLIQFLGDAIVLPAGALHQVNFHSCIQVTFDFVSPEHLVESFHLTQELR 2120
 662 QY 662 SLPP-----DHRMLYAQMRAVFAQVAAVAGALQ 690
 2121 DB 2121 LLKEEINYDDKL---QVKNILYHAVKENVRAK 2150

RESULT 12
 Q5RCM5_PONPY PRELIMINARY; PRT; 1441 AA.
 ID Q5RCM5_PONPY PRELIMINARY; PRT; 1441 AA.
 AC Q5RCM5_PONPY PRELIMINARY; PRT; 1441 AA.
 DT 01-FEB-2005 (TREMBLrel. 29, Created)
 DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
 DE Hypothetical protein DKFP469A034 (Fragment).
 GN Name=DKFP469A034;
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Pongo.
 ON NCBI_TaxID=9600;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC The German cDNA Consortium;
 RG Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
 RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL, CR858245; CAH90482.1; -, mRNA.
 DR InterPro; IPR003347; TF_JmJC.
 DR Pfam; PF02373; JmJC; 1.
 DR SMART; SM00558; JmJC; 1.
 KW Hypothetical protein.
 FT NON TER
 SQ SEQUENCE 1441 AA; 160281 MW; 70499305EACCI177C CRC64;

Query Match 19.4%; Score 719.5; DB 2; Length 1441;
 Best Local Similarity 29.4%; Pred. No. 9.6e-39;
 Matches 221; Conservative 115; Mismatches 270; Indels 145; Gaps 25;

61 QY 61 AKHLLSGLDRLCLRLKREERAAWAQREGQGPAMTEDSPGIPHCCSRCHHGLFNTHWRC 120
 703 DB 703 SKYILDIIGDKFCQLVTSKALSVVKKDAK-IANKRAVRGVLEMCACETLFIHWHVC 761
 121 QY 121 SHCSHRLCVACRGIAGAKNREKTSQSOQHT-DDCAQEAHGAACSLILTOFVSSQALAE 179
 762 DB 762 QKCGFVCLDCYK-----AKERKSRDKELYAMKCVKGQPHDHHKHLMTQIIFGSVLTDL 817
 180 QY 180 STVMHQAQWAKFDIRGHCF-----QVDARVWAPGDGQOQKEPTEKTP 221
 818 DB 818 LDAMHTLEKYGKSHCHCTKNQLQVGNFPTMNGVSVQLVNLHNSKLSLCPESQOQ 877
 222 QY 222 PTPQPS-NGQSNRTKQIKER---TPPOSTESP----- 249
 878 DB 878 NTPQKSEKNGGSSPESDVGTDNKLTPPESQSLPHLADLAQKAREBKENKELTLENQI 937
 250 QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCGHDIRHMAFAPV-----T 290
 938 DB 938 KEERQDNSESPNGKTSPLVSONNEQGSGLRDLTLTTAGKLRVGSTDAGTAFAPVSMGA 997
 291 QY 291 PALPDDDRITNILDITIAQVVERKI-----QBKALGPGLAGAGSLRKG 333
 998 DB 998 PSSKSGRTWPNILDDITASVVENKIPSKTSKINVKPELKEEPESITSAVDENNKLYS- 1056
 334 QY 334 LSLPLSVRTRLSPGALLWLTQBRPRKHGFLFOHWRQOQPVLVSGIQKTLRLSLWGM 393

1057 -DIPHSWICEK-----HILWLKDYKNSNNWKLFKECWKQGOQAPVVGHVKKNNISLWKA 1111
 394 QY 394 ALG-TLGGQVOSLTALGPPQPTNLDSTAFWEGFSGHPETRPKLDGSGVLLHRLTGLDKD-- 450
 1111 DB 1111 SISLDFGDHQADLLNCKDSIISNANVKEFWDFEVSCKOKNKGSETVVLKL-----KDW 1166
 451 QY 451 -----ASRVQNLVSSLPPEYCAHQGKLNLSYLPGLTLHP-LEPOLMAAYGVNS 500
 1167 DB 1167 SGEDFKTWPARYEDLLKSLPLPEYCNPEKGFNLASHLP-GFFVRPDLGPRLCSAYGVA 1225
 501 QY 501 HRGH-LGTKNLCVEVSDILISLVHAEQAQLPPWYAKQDFLSGLDGEGLWS----- 549
 1226 DB 1226 AKDHGIGITNLHIEVSDVNVILVYGTAKNGIILSKAGILKKFEEEDDDILRKRLKDD 1285
 550 QY 550 --PGSQTSTVHVHFRADQAOIRRRFLQVMVC-PAGAGTL---EPGAPGSCYLDAGLRRRL 603
 1286 DB 1286 EIPGA----LRHIYAGKVDVKIREFLQKISKEQGLEVLPEHDPIRDQSWYVNNKLRQL 1341
 604 QY 604 EEMGVSWTLLQAPGEAVLPACAPHOVQGLVSTISVTOHFLSPETSALSQAOLCHQGA 663
 1342 DB 1342 BEYGVRTCTLIQFLGDAIVLPAGALHQVNFHSCIQVTFDFVSPEHLVESFHLTQELRL 1401
 664 QY 664 PP-----DHRMLYAQMRAVFAQVAAVAGALQ 690
 1402 DB 1402 KEEINYDDKL---QVKNILYHAVKENVRAK 1429

RESULT 13
 Q7Z3E7_HUMAN PRELIMINARY; PRT; 2210 AA.
 ID Q7Z3E7_HUMAN PRELIMINARY; PRT; 2210 AA.
 AC Q7Z3E7;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-NAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Hypothetical protein DKFP686K091.
 GN Name=DKFP686K091;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 ON NCBI_TaxID=9606;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Human cervix;
 RA Pousta A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
 RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BX537954; CAD97921.1; -, mRNA.
 DR Ensembl; ENSG00000171988; Homo sapiens.
 DR InterPro; IPR003347; TF_JmJC.
 DR Pfam; PF02373; JmJC; 1.
 DR SMART; SM00558; JmJC; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 2210 AA; 245646 MW; B906415F74BBD690 CRC64;

Query Match 19.4%; Score 719; DB 2; Length 2210;
 Best Local Similarity 30.0%; Pred. No. 1.8e-38;
 Matches 207; Conservative 108; Mismatches 250; Indels 124; Gaps 23;

61 QY 61 AKHLLSGLDRLCLRLKREERAAWAQREGQGPAMTEDSPGIPHCCSRCHHGLFNTHWRC 120
 1515 DB 1515 SKYILDIIGDKFCQLVTSKALSVVKKDAK-IANKRAVRGVLEMCACETLFIHWHVC 1573
 121 QY 121 SHCSHRLCVACRGIAGAKNREKTSQSOQHT-DDCAQEAHGAACSLILTOFVSSQALAE 179
 1574 DB 1574 QKCGFVCLDCYK-----AKERKSRDKELYAMKCVKGQPHDHHKHLMTQIIFGSVLTDL 1629
 180 QY 180 STVMHQAQWAKFDIRGHCFQVDARVW-----PGDQGOQKEPTEKTPPTPQSCNGSDSNRTK 236
 1630 DB 1630 LDAMHTLEKYGKSHCHCTKNQLQVGNFPTMNGVSVQSQOQTPPKSEK--NGGSSPES 1687
 237 QY 237 DIKEE---TPPOSTESP-----AEDGAGR- 256

```
Db 1688 DVGTDNKLTTPESQSPHLWLADLAEQKARBEKENKELTLENQIKEEREQDNSESPNGRT 1747
QY 257 SPLPC-----PSICELLASTAVKLCGLGHDRIHMAFAPV-----TPALPSDDRTITNILD 306
Db 1748 SPLVSONNEOGSLRDLTLTTAGKLRVGTSDAGIAFAPVYMGAPSKSGRTWPNILDDI 1807
QY 307 IAOVVERKI-----QEKALGFLRAGSLRGKLSLPLSPVTRTRLSPPG 349
Db 1808 IASVVENKIPPSKTSKINVKPELKEPEESIIISAVDENNNKLYS-DIPHSWICEK----- 1860
QY 350 ALLWLQEPKPHGFHLFOEHWROGQPVLSGQIKTLRLSLWGMALG-TLGGQVQSILAL 408
Db 1861 HILWLKDYKNSSNNKLFKECWKGQPAVSGVHKNNISLWKAESISLDFGDHQADLLNC 1920
QY 409 GPPOPTNLDTAFWEGFHPETRPKLDGSLVLLHRLTLGDKD-----ASRVQNL 457
Db 1921 KDSIISNANVKEFWDFEESVKQKNSGETVVLKL-----KDPGSGEDFKTWPMPARVEDL 1976
QY 458 VSSLPLPEYCAHQGKUNLASYLPLGLTLHP-LEPQLWAAVGVNSHRGH-LGTKNLCVETS 515
Db 1977 LKSLPLPEYCNPEGKFNLAGHLP-GFFVRPDLGRLCSAYGVVAAKDHDIGTTNLHIEVS 2035
QY 516 DLISILVHAAQALPPWYRAQKDFLSGLDGEGLWS-----PGSQTSTVWHVFA 563
Db 2036 DVYNILVYVGIAGKNGTLKAGILKKFEEDLDILRLKRLKDSSEIPGA----LWHIYAG 2091
QY 564 QDAQIRRRFLQWVC-PAGAGTL---EPGAPGSCYLDAGRLRLRREEMGVSCWTLLOAPGE 619
Db 2092 KDVKIKREFLQKTSKEQGLEVLPEHDPIRQDSWYVNNKLRQLLEEYGVRTCTLIQLGLD 2151
QY 620 AVILVPAGAHQVQGLVSTISVTQHFLSPE 648
Db 2152 AIVLPAGALHQVQNFHSCIQVTEDFVSPE 2180

RESULT 14
Q692K6_MOUSE Q692K6_MOUSE PRELIMINARY; PRT; 2428 AA.
AC Q692K6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MKIAA1380 protein (Fragment).
GN Name=JmJdic; Synonyms=mkIAA1380;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryonic intestinal tract;
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:205-218(2004).
DR EMBL; AKI73162; BAD32440.1; -; mRNA.
DR MGI; MGI:1918614; JmJdic.
DR InterPro; IPR003347; TF_JmJc.
DR Pfam; PF02373; JmJc; 1.
DR SMART; SM00558; JmJc; 1.
FT NON TER 1
SQ SEQUENCE 2428 AA; 269552 MW; D673669EDDE67779 CRC64;

Query Match 19.0%; Score 702; DB 2; Length 2428;
Best Local Similarity 28.7%; Pred. No. 2.6e-37;
Matches 215; Conservative 122; Mismatches 270; Indels 142; Gaps 28;

QY 61 AKHLLSGDLRLCLRLKREALAWAQEGQGPANTEDSGIPHCSCRHGLFNTHWRC 120
```

```
Db 1691 SKYILDIIGDFKQLVTSEKLTALSMWYKDAK-IAWKRAVRGVREMDACETATFNHVWC 1749
QY 121 SHCSHRLCVACGRIAGAGKNREKTSQEOHT-DDCAQEAAGHAACSLILTLQFVSSQALAE 179
Db 1750 RKCGFVACLDYK----AKERSSRDKELYAMWKCVKGQPHDKHLMPTQIIIFGSVLTDL 1805
QY 180 STVMHQAWAKFDIRGHCFC-----QV-----DARVWAPGQGG 212
Db 1806 LDAMHILREKYGKISHCHCTNRQNLQGGNVPTMNGVSQVLQNLVHHSNKTSLP-ESQ 1864
QY 213 QKEPTEK-----TPPTQPSC-----NGDSNRRTKDI 238
Db 1865 QNSPQSQTNNGNSSPSGASDLSLTPTPESQSPHLWLADLAQKREKEQENKFTLEREI 1924
QY 239 KEB-TPDSTESP---AEDGAGRSPLPCPSLCELLASTAVKLCGLGHDRIHMAFAPV----T 290
Db 1925 KEGDQDASDSPNGSTSPASQSQNEQSTLRDLTTTAGKLRVGTSDAGIAFAPVYMG 1984
QY 291 PALPSDDRTITNILDIIAQVVERKIQEKALGPGLRAGSGLR-----KGLSLPLSPV-- 341
Db 1985 SSCKGGRTPNILDIIASVVENKIP-----PNKTSKINIKSEPNBEPKESLIPATDES 2039
QY 342 RTLSPPGA-----LLWLQEPKPHGFHLFOEHWROGQPVLSGQIKTLRLSLWGMAL 395
Db 2040 KSYRDIPIHSHWICDQHLMLKDYKNNNWKLKFCWKQKGQPAVSGVHKNNISLWKAESI 2099
QY 396 G-TLGGVQSLTALGPQPPTNLDTAFWEGFHPETRPKLDGSLVLLHRLTLGDKDA--- 451
Db 2100 SLDFGDHQADLLNCKDSISVSNANVKEFWDFEESVKQKNGGETVVLKL---KDCPSG 2155
QY 452 -----SRVQNLVSSLPLPEYCAHQGKUNLASYLPLGLTLHP-LEPQLWAAVGVNSHR 502
Db 2156 EDFKAMPTRYEDFLRCLPLPEYCNPEGKFNLAGHLP-GFFVRPDLGRLCSAYGVVAAK 2214
QY 503 GH-LGTKNLCVEVSDILSILVH-----AEAQLPPMYRAQK--DFLSGLDGEGLWS 549
Db 2215 DHDIGTTLNLHIEASDVVNVVLYVYVGIAGKNGVLKAGILKKFEEDLDILRLKRLKDSSEI 2274
QY 550 PGSQTSVWHVFRQAQRIIRRELQWVC-PAGAGTL---EPGAPGSCYLDAGLRRRLREE 605
Db 2275 PGA----LWHIYAGKVDKIREFLQKTSKEQGLEVLPEHDPIRQDSWYVNNKLRQLLEE 2330
QY 606 WGVSCWTLLOAPGEAVLVPAGAHQVQGLVSTISVTQHFLSPSPTSALSALQCHQASLPP 665
Db 2331 YGVRACTLIQLGDAIVLPAGTLHQVQNFHSCVQVTEDFVSPEHLVQSPHLTQELRLKE 2390
QY 666 ----DHRMLYAQMDRAVFOAKAAGALQ 690
Db 2391 EINYYDKL----QVKNILYHAVKEMVRALK 2416

RESULT 15
Q9NYP4_HUMAN Q9NYP4_HUMAN PRELIMINARY; PRT; 1417 AA.
AC Q9NYP4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative zinc finger protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Bone marrow;
RX MEDLINE=20541717; PubMed=11087669; DOI=10.1006/geno.2000.6345;
RA Lai F., Godley L.A., Fernald A.A., Orelli B.J., Pamintuan L., Zhao N.,
RA Le Beau M.M.;
RT "cDNA cloning and genomic structure of three genes localized to human
RT chromosome band 5q31 encoding potential nuclear proteins.";
RT
```


Search completed: November 25, 2005, 22:10:39
Job time : 240 secs

```

RL Genomics 70:123-130(2000).
DR EMBL; AF251039; AAF63765.1; -; mRNA.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR003347; TF_JmJC.
DR Pfam; PF02373; JmJC; 1.
DR SMART; SM00558; JmJC; 1.
SQ SEQUENCE 1417 AA; 154720 MW; E9CF503D8654FD2B CRC64;

Query Match      18.9%; Score 700.5; DB 2; Length 1417;
Best Local Similarity 26.7%; Pred. No. 1.7e-37;
Matches 237; Conservative 120; Mismatches 291; Indels 239; Gaps 32;

QY 6 SCVQAAGEVGLTGHQKRRSPLEE-----KQLEEDSSATSE----- 44
Db 559 SCINVA-----PHLHKCRECLERYRKFKQEODDSTVACRFFHFRRLIFTRKGLRV 611
QY 45 EGGGGPG-----PEASLNGU-----AKHLLSGLDRLCLRLKEREALAWAORE 89
Db 612 EGLSPQSDPDAMNLIWIPSSLAEGIDLETSKYIILANVGDFCQLVNSEKEAMVMVEPH 671
QY 90 GQGPAMTEDSPGIPCHCCHGLFNTHWRCSHCHRLCVACGRIAGAGKREKTGSOEQ 149
Db 672 -QXVANKRAVRGVREWCDCVETTLFNIHWVCRKCGFGVCLDCYRLR---KSRPRSETEM 727
QY 150 HTDD-----CAQEAAGAACSLILTFQVSSQALAEISTVMHQAWAKEDIRGHCF- 201
Db 728 GDEEVFWLKCAGQSHPEMLPTQIIPGTALYNIQDMVHAARGKWGIKANCPCISRON 787
QY 202 ARWAP-----GDGQOQKEPTEKPTPTPOPSCNGDSNRTK 236
Db 788 KSVLRPAVNTGMSQLPSINPSASSGNETTFSGGG-----PAPVTTPEPDHVPKADST--- 840
QY 237 DIKEETPDSTESPAEDGAGRSPL-----PCP-----SLCELLASTAVK----- 274
Db 841 DIRSEPLKTDSSASH--SNSSELKAIKPPCPDTPAPSSALHMLADLADATQAKAKEETKEAGS 898
QY 275 -----LCLG-----HDRIH----- 283
Db 899 LRSVLNKESHSPGLDSFNSTAKVSLTPKLFNSLLGLPTASNNTKTEGSSLRDLLHSGPG 958
QY 284 -----MAPAPV-----TPALPSDDRIITNILDSIIAQVVERKIOEKA-----LGPG 323
Db 959 KLPQTLDTGIPTFPVPVSTSSAGVSKASLPNPLDHIASVVENKKTSDASKRACNLDT 1018
QY 324 LRAGSLRKGLSLPLSPVTRLSP-----PGALLWLOEPRPKHGFHLFOEHWFOGQP 375
Db 1019 QKEVKEMVNLV-----LDPHTSHSWLDCGRLLCLHDPSNKNKNWKIFRECWKQGP 1070
QY 376 VLVSGTQKTLRLSLMGMEALGTIGG--VQOSLTALGPPQPTNLDSTAFWEGFESHPETRPK 433
Db 1071 VLVSGVHKKLKSELWKPEAFSQEFQDQDVLVNCRCNAIISDVKVRDFWDGFEIICKRLR 1130
QY 434 LDEGSVLLH-----RTLGDKASRVONLVSSPLPEYCAHQKLNLASVLPGLTLH 486
Db 1131 SEDGQPMVLKLDWPPGEDFRDMPTFRFEDLMENLPLPEYTKRDLNLASRLPSYFVRP 1190
QY 487 PLEPOLWAAYG-VNSHRHLGTNLCVEUSDLSILVHAEALPPMYRA-QKDFLSGLD- 543
Db 1191 DLGPKMYNAYGLITAEDRVGTNTLHLDVSDAVNMVY--VGIPIGEGAHDDEVLTIDE 1248
QY 544 -----GEGLSWPGSQSTVMHVFRAODAOIRRFLOWVCPAGAGTLEPGAGP-- 590
Db 1249 GDADEVTKQRIHDGKEKPA-----LWHIYAADAEKIRELLRKV---GEEQOENPPDHD 1301
QY 591 -----SCYLDAGLRLRLREWGVCWTLTQAFPEAVLVPAGAPHQVQGLVSTISVTQHFL 645
Db 1302 PTHDQSWYLDQTLRKELYEYGVQGWAIYQFLGDVFIIPAGAPHQVHNLVYSCIKVAEDFV 1361
QY 646 SPETSALSQQLCHQASLPPDH--RMLYAQMDRAVFOAVKAAGALQ 690
Db 1362 SPEVHKCFRLTQEPFHLSTNTHNEDKLQVKNIIYHAVKDAVGTILK 1408

```

This Page Empty (10-10)